

70884

**STIC-Biot ch/ChemLib**

Fr m: Spector, Lorraine  
Sent: Monday, July 15, 2002 1:52 PM  
To: STIC-Biotech/ChemLib  
Subject: Search request, 09/202054

Please TRANSLATE SEQ ID NO:2, and then do a word/fragment search with a word size of 6 aa.  
Also search SEQ ID NO:1, residues 30-836, word/fragment search with a word size of 6 aa.  
Need Pending, Issued, and Commercial databases for both.

Thanks.

Lorraine Spector  
Patent Examiner  
Art Unit 1647  
CM1 10B-11  
703-308-1793

**Point of Contact:**  
**Toby Port**  
**Technical Info. Specialist**  
**CM1 6A04**  
**703-308-3534**

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 7/16  
Date Completed: 7/17  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

**TYPE OF SEARCH:**  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

**VENDOR/COST (where applic.)**  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THE BUREAU OF THE

POST OFFICE

WASHINGTON, D. C.

1877-1878

1877-1878



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2002, 09:44:08 ; Search time 37.52 Seconds  
(without alignments)  
2389.034 Million cell updates/sec

Title: US-09-202-054-1\_COPY\_30\_836  
Perfect score: 807  
Sequence: 1 FPKTLPDVTLDVVKNHVIV.....PGAHKQGSVISLDLYTCELD 807

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues  
Word size : 6

Total number of hits satisfying chosen parameters: 4655

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : A\_Geneseq\_032802:\*  
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	807	100.0	1045	20	Human DNAX toll-11
2	807	100.0	1049	20	Human PRO285 prote
3	807	100.0	1049	20	Human Toll protein
4	807	100.0	1049	21	Human PRO285 prote
5	807	100.0	1049	22	Human brain expres
6	404	50.1	432	22	Human immune/haema
7	57	7.1	121	22	Human signal trans
8	41	5.1	161	22	Novel signal trans
9	41	5.1	161	22	Human polypeptide
10	41	5.1	235	22	Human polypeptide
11	34	4.2	394	20	Mouse DNAX toll-11

100	2.0	16	100	20	AAW86362	Mouse DNAX toll-11
12	1.5	12	483	22	AAG65892	Amino acid sequenc
13	1.5	12	605	22	AAG65893	Amino acid sequenc
14	1.4	11	14	22	AAE07271	Human leucine-rich
15	1.4	11	14	22	AAE07271	Human leucine-rich
16	1.4	11	14	22	AAE06918	Human Garp protein
17	1.4	11	14	22	AAE07278	Human leucine-rich
18	1.4	11	14	22	AAE07281	Human leucine-rich
19	1.4	11	14	22	AAU29303	Human PRO polypept
20	1.4	11	14	22	AAE07266	Human leucine-rich
21	1.4	11	14	22	AAE07280	Human leucine-rich
22	1.2	10	329	20	AAW86356	Partial human DNAX
23	1.2	10	426	22	AAW864892	Human secreted pro
24	1.2	10	504	22	AAU14800	Novel bone marrow
25	1.2	10	1040	22	ABB23027	Protein #5026 enco
26	1.2	10	1040	22	AAW31141	Peptide #5178 enco
27	1.2	10	1041	20	AAW41768	Human PRO286 prote
28	1.2	10	1041	20	AAW05867	Human Toll protein
29	1.2	10	1041	21	AAW44324	Human PRO286 prote
30	1.1	9	15	12	AAW13130	GP1b alpha peptide
31	1.1	9	114	22	AAW40831	Human polypeptide
32	1.1	9	155	21	AAW44116	Human cancer assoc
33	1.1	9	158	21	AAW94900	Human secreted pro
34	1.1	9	227	22	AAW64951	Human secreted pro
35	1.1	9	234	14	AAW42265	Decorin sequence p
36	1.1	9	235	22	AAU20516	Human secreted pro
37	1.1	9	280	14	AAW42266	Decorin sequence p
38	1.1	9	293	10	AAW91368	45 kDa amino termi
39	1.1	9	305	14	AAW42267	Decorin sequence p
40	1.1	9	331	14	AAW42260	Mature decorin PT-
41	1.1	9	342	17	AAW89439	Human recombinant
42	1.1	9	359	21	AAW57079	Human decorin amin
43	1.1	9	359	22	AAW78511	Human decorin amin
44	1.1	9	493	22	AAW49650	Human SEC2 protein
45	1.1	9	557	22	AAW39045	Human polypeptide
46	1.1	9	562	21	AAW49933	Human glycoprotein
47	1.1	9	568	21	AAW49935	Human glycoprotein
48	1.1	9	610	15	AAW51116	Platelet glycoprot
49	1.1	9	610	17	AAW56664	Mutated platelet gl
50	1.1	9	610	17	AAW89436	Platelet platelet g
51	1.1	9	610	18	AAW18201	Platelet glycoprot
52	1.1	9	626	22	AAE12135	Human glycoprotein
53	1.1	9	799	20	AAW86352	Human DNAX toll-11
54	1.1	9	837	20	AAW86361	Human DNAX toll-11
55	1.1	9	953	22	ABB63934	Drosophila melanog
56	1.1	9	1315	22	ABB58454	Drosophila melanog
57	1.1	9	1388	17	AAW89471	Collagen/decorin f
58	1.1	9	1388	21	AAW84539	Amino acid sequenc
59	1.0	8	14	22	AAE07267	Human leucine-rich
60	1.0	8	14	22	AAE07270	Human leucine-rich
61	1.0	8	105	21	AAW27754	Protein fragment e
62	1.0	8	134	22	AAW23966	Human EST encoded
63	1.0	8	149	21	AAW42453	Human OREP ORF2217
64	1.0	8	162	21	AAW25207	Eucalyptus grandis
65	1.0	8	165	21	AAW23658	Arabidopsis thalia
66	1.0	8	275	21	AAW05586	Arabidopsis thalia
67	1.0	8	305	22	AAW65805	RPV5-like protein
68	1.0	8	344	22	ABB69226	O. longistaminata
69	1.0	8	371	21	AAW05585	Arabidopsis thalia
70	1.0	8	393	21	AAW05584	Arabidopsis thalia
71	1.0	8	408	22	AAW03522	Human polypeptide
72	1.0	8	428	20	AAW34579	Porphorymonas ging
73	1.0	8	437	20	AAW34436	Porphorymonas ging
74	1.0	8	579	22	AAW65802	RPV5-like protein
75	1.0	8	611	20	AAW93594	O. longistaminata
76	1.0	8	636	21	AAW32218	Arabidopsis thalia
77	1.0	8	735	21	AAW32217	Arabidopsis thalia
78	1.0	8	743	21	AAW32216	Arabidopsis thalia
79	1.0	8	796	21	AAW88054	Human mddt protein
80	1.0	8	806	17	AAW09254	Tomato pathogen re
81	1.0	8	806	17	AAW09254	Tomato pathogen re
82	1.0	8	806	17	AAW85301	Tomato pathogen re
83	1.0	8	863	16	AAW75919	Tomato Cf-9. Lyc
84	1.0	8	863	17	AAW85296	Tomato pathogen re

85	1.0	895	17	AAR91615	Trypanosoma cruzi	158	7	0.9	195	18	AAW44801	B. burgdorferi str
86	1.0	928	22	AAB95528	Human protein sequ	159	7	0.9	195	21	AAW84723	A decorin binding
87	1.0	1024	17	AAW03184	Rice Xa21 (RRK-F)	160	7	0.9	198	21	AAW25125	Pinus radiata cell
88	1.0	1025	17	AAW03185	Rice Xa21 disease	161	7	0.9	202	22	ABG28119	Novel human diago
89	1.0	1025	20	AAW93597	O. longistaminata	162	7	0.9	204	20	AAW60512	Human normal blad
90	1.0	1025	22	AAW67442	Amino acid sequenc	163	7	0.9	204	20	AAW34533	Porphyrmonas ging
91	1.0	1026	20	AAW39279	Xa21 signal transd	164	7	0.9	205	21	AAW25288	Eucalyptus grandis
92	1.0	1176	22	AAW79269	Human protein SQ	165	7	0.9	206	18	AAW14574	Streptococcus pneu
93	1.0	1176	22	AAW70151	DNA encoding human	166	7	0.9	207	22	ABW68307	Drosophila melanog
94	1.0	1216	22	ABW60349	Drosophila melanog	167	7	0.9	211	21	AAW25443	Pinus radiata cell
95	0.9	14	22	AAE07269	Human leucine-rich	168	7	0.9	212	22	ABW17119	Human nervous syst
96	0.9	35	22	ABW37647	Peptide #5153 enco	169	7	0.9	216	22	AAU29932	Novel human secret
97	0.9	35	22	AAW18595	Peptide #5029 enco	170	7	0.9	221	20	AAW34400	Porphyrmonas ging
98	0.9	38	21	AAW15347	Arabidopsis thalia	171	7	0.9	221	22	AAW93709	Human protein sequ
99	0.9	38	21	AAW48691	Arabidopsis thalia	172	7	0.9	224	22	AAW70764	S cerevisiae apopt
100	0.9	39	21	AAW04312	Arabidopsis thalia	173	7	0.9	226	22	ABG15835	Novel human diago
101	0.9	41	22	AAW85849	Human immune/haema	174	7	0.9	228	21	AAW57394	Arabidopsis thalia
102	0.9	43	19	AAW75064	Human secreted pro	175	7	0.9	237	21	AAW19183	Lipid metabolism p
103	0.9	48	22	ABW30631	Peptide #3282 enco	176	7	0.9	239	21	AAW09390	Arabidopsis thalia
104	0.9	48	22	ABW35799	Peptide #3305 enco	177	7	0.9	241	21	AAW25748	Arabidopsis thalia
105	0.9	48	22	ABW21218	Protein #3217 enco	178	7	0.9	241	22	ABG26463	Novel human diago
106	0.9	48	22	AAW56607	Human brain expres	179	7	0.9	253	22	AAW23550	Human EST encoded
107	0.9	48	22	AAW68985	Human bone marrow	180	7	0.9	256	22	AAW65886	Amino acid sequenc
108	0.9	48	22	AAW16815	Peptide #3249 enco	181	7	0.9	257	21	AAW25424	Pinus radiata cell
109	0.9	48	22	AAW29299	Peptide #3336 enco	182	7	0.9	266	21	AAW41583	Human OREF ORF1347
110	0.9	48	22	AAW04531	Peptide #3213 enco	183	7	0.9	269	21	AAW09389	Arabidopsis thalia
111	0.9	62	22	AAW87747	Human immune/haema	184	7	0.9	269	21	AAW59238	Arabidopsis thalia
112	0.9	74	22	AAW08330	Human polypeptide	185	7	0.9	269	21	AAW59520	Arabidopsis thalia
113	0.9	94	20	AAW78227	Fragment of human	186	7	0.9	270	21	AAW57551	Arabidopsis thalia
114	0.9	96	21	AAW23934	Arabidopsis thalia	187	7	0.9	271	21	AAW22003	Arabidopsis thalia
115	0.9	97	21	AAW25344	Pinus radiata cell	188	7	0.9	292	22	AAW91002	Arabidopsis thalia
116	0.9	102	21	AAW13274	Caenorhabditis ele	189	7	0.9	292	22	AAE03524	Human secreted pro
117	0.9	102	21	AAW07741	Human polypeptide	190	7	0.9	294	21	AAW25746	Arabidopsis thalia
118	0.9	103	21	AAW41538	Human OREF ORF1302	191	7	0.9	302	22	AAW96145	Arabidopsis thalia
119	0.9	108	22	AAW90765	Human immune/haema	192	7	0.9	306	21	AAW57393	Putative P. abyssi
120	0.9	115	22	AAW87595	Human secreted pro	193	7	0.9	307	22	AAW38910	Human polypeptide
121	0.9	123	20	AAW35455	Human haematologic	194	7	0.9	311	21	AAW57392	Arabidopsis thalia
122	0.9	123	22	AAW95966	Human haematologic	195	7	0.9	322	22	AAU51780	Propionibacterium
123	0.9	123	22	AAW11458	Human reproductive	196	7	0.9	325	21	AAW26324	Arabidopsis thalia
124	0.9	126	22	AAW11174	Human polypeptide	197	7	0.9	325	21	AAW48081	Arabidopsis thalia
125	0.9	146	21	AAW31174	Novel human secret	198	7	0.9	326	22	ABW61863	Drosophila melanog
126	0.9	146	21	AAW25197	Eucalyptus grandis	199	7	0.9	327	14	AAW34507	T.aquaticus malate
127	0.9	146	21	AAW01595	Human secreted pro	200	7	0.9	328	21	AAW26323	Arabidopsis thalia
128	0.9	146	22	AAW81086	Human haematologic	201	7	0.9	329	21	AAW48080	Arabidopsis thalia
129	0.9	150	21	AAW02236	Human secreted pro	202	7	0.9	329	21	AAW22044	Arabidopsis thalia
130	0.9	154	21	AAW25131	Pinus radiata cell	203	7	0.9	331	21	AAW37413	Arabidopsis thalia
131	0.9	154	21	AAW25449	Pinus radiata cell	204	7	0.9	332	21	AAW58898	Breast and ovarian
132	0.9	157	21	AAW25122	Eucalyptus grandis	205	7	0.9	334	18	AAW19209	Human interleukin-
133	0.9	157	22	AAU42009	Propionibacterium	206	7	0.9	334	21	AAW20861	Arabidopsis thalia
134	0.9	160	20	AAW82488	Ehrlichia sp. E74	207	7	0.9	339	21	AAW25524	Pinus radiata cell
135	0.9	172	19	AAW77625	DNA-invertase Hin	208	7	0.9	342	21	AAW20860	Arabidopsis thalia
136	0.9	172	19	AAW48618	Staphylococcus aur	209	7	0.9	343	21	AAW25517	Pinus radiata cell
137	0.9	173	21	AAW25395	Pinus radiata cell	210	7	0.9	347	21	AAW09388	Arabidopsis thalia
138	0.9	173	21	AAW74996	Neisseria gonorrhe	211	7	0.9	347	21	AAW59237	Arabidopsis thalia
139	0.9	173	21	AAW74997	Neisseria meningit	212	7	0.9	347	21	AAW59241	Arabidopsis thalia
140	0.9	173	21	AAW74998	Neisseria meningit	213	7	0.9	347	21	AAW59519	Arabidopsis thalia
141	0.9	177	22	AAU59051	Propionibacterium	214	7	0.9	348	22	ABG05182	Novel human diago
142	0.9	178	22	AAW93380	Human polypeptide	215	7	0.9	352	21	AAW09955	Arabidopsis thalia
143	0.9	179	22	ABW66444	Drosophila melanog	216	7	0.9	352	21	AAW59240	Arabidopsis thalia
144	0.9	180	21	AAW25324	Eucalyptus grandis	217	7	0.9	352	21	AAW59518	Arabidopsis thalia
145	0.9	183	22	ABW10514	Human cDNA SQ ID	218	7	0.9	353	21	AAW24026	Human PRO1788 prot
146	0.9	183	22	AAU18041	Human immunoglobul	219	7	0.9	353	21	AAW99455	Human PRO1788 (UNQ
147	0.9	184	22	AAW15207	Human colon cancer	220	7	0.9	353	22	AAE03427	Human gene 1 encod
148	0.9	185	18	AAW14566	Streptococcus pneu	221	7	0.9	353	22	AAE03459	Human gene 1 encod
149	0.9	185	21	AAW25198	Eucalyptus grandis	222	7	0.9	353	22	AAU00472	Human TANGO 366 pr
150	0.9	189	21	AAW14569	Streptococcus pneu	223	7	0.9	353	22	AAW88413	Human membrane or
151	0.9	190	18	AAW14569	Streptococcus pneu	224	7	0.9	353	22	AAW66204	Protein of the inv
152	0.9	191	20	AAW32126	Human LYST interac	225	7	0.9				
153	0.9	192	19	AAW65023	Tsuga plicata diri	226	7	0.9				
154	0.9	192	22	AAE06536	Arabidopsis thalia	227	7	0.9				
155	0.9	194	17	AAW07185	B. burgdorferi HB-	228	7	0.9				
156	0.9	194	18	AAW44799	B. burgdorferi str	229	7	0.9				
157	0.9	195	18	AAW44796	B. burgdorferi str	230	7	0.9				

231	7	0.9	355	22	AAB31958	Amino acid sequenc	304	7	0.9	546	22	AAB80254	Human PRO247 prote
232	7	0.9	357	21	AAB58365	Lung cancer associ	305	7	0.9	549	21	AAG35762	Arabidopsis thalia
233	7	0.9	357	21	AAB40712	Human ORFX ORF476	306	7	0.9	550	21	ABB67024	Drosophila melanog
234	7	0.9	358	21	AAG09954	Arabidopsis thalia	307	7	0.9	557	21	AAG53564	Arabidopsis thalia
235	7	0.9	358	21	AAG59239	Arabidopsis thalia	308	7	0.9	565	22	AAM41692	Human polypeptide
236	7	0.9	360	21	AAG09953	Arabidopsis thalia	309	7	0.9	575	20	AAW83389	Caenorhabditis ele
237	7	0.9	360	22	AAG91578	C glutamic prote	310	7	0.9	577	22	ABG23262	Novel human diagno
238	7	0.9	360	22	AAU20546	Human secreted pro	311	7	0.9	593	21	AAAY44945	Wheat sulphate per
239	7	0.9	362	22	AAM93819	Human polypeptide,	312	7	0.9	593	22	AAW79839	Human protein SEQ
240	7	0.9	362	22	AAW38909	Human polypeptide,	313	7	0.9	593	22	AAW79839	Human protein SEQ
241	7	0.9	362	22	AAW38909	Human polypeptide,	314	7	0.9	598	21	AAG20654	Arabidopsis thalia
242	7	0.9	362	22	AAW38909	Human polypeptide,	315	7	0.9	599	21	AAG20653	Arabidopsis thalia
243	7	0.9	363	22	ABW70103	Drosophila melanog	316	7	0.9	613	22	AAW82888	Maize fasciated ea
244	7	0.9	370	20	AAW56108	Human int-1 mammar	317	7	0.9	616	21	AAW99347	Human PRO1113 (UNQ
245	7	0.9	370	21	AAW19786	Human Wnt-1 protei	318	7	0.9	616	22	AAU29102	Human PRO polypept
246	7	0.9	370	21	AAW19786	Human Wnt-1 protei	319	7	0.9	616	22	AAW66096	Protein of the inv
247	7	0.9	370	21	AAW19786	Human Wnt-1 protei	320	7	0.9	623	21	AAW25493	Eucalyptus grandis
248	7	0.9	371	21	AAW56812	Human prostate can	321	7	0.9	627	19	AAW62451	Mycoplasma hyopneu
249	7	0.9	371	22	ABG13003	Novel human diagno	322	7	0.9	631	21	AAG25781	Arabidopsis thalia
250	7	0.9	371	22	AAW40695	Human polypeptide	323	7	0.9	632	21	AAG25780	Arabidopsis thalia
251	7	0.9	371	22	AAW40695	Human polypeptide	324	7	0.9	638	21	AAG25779	Arabidopsis thalia
252	7	0.9	374	18	AAW31309	Human polypeptide	325	7	0.9	642	20	AAW13396	Amino acid sequenc
253	7	0.9	386	21	AAW59236	Arabidopsis thalia	326	7	0.9	642	21	AAW33425	Human PRO332 prote
254	7	0.9	394	22	ABB64698	Drosophila melanog	327	7	0.9	642	22	AAU12356	Human PRO332 polyp
255	7	0.9	396	19	AAW62598	Human 7-transmembr	328	7	0.9	642	22	AAW80264	Human PRO332 prote
256	7	0.9	396	22	AAW77798	Human EX33 G-prote	329	7	0.9	643	22	AAW65906	Amino acid sequenc
257	7	0.9	402	21	AAW29442	Arabidopsis thalia	330	7	0.9	650	21	AAW82527	Human RP105 protei
258	7	0.9	404	22	AAW99936	Human expressed po	331	7	0.9	653	22	AAW65907	Amino acid sequenc
259	7	0.9	409	21	AAW26439	Drosophila melanog	332	7	0.9	659	17	AAW89265	Heparinase-III, F
260	7	0.9	410	20	AAW37460	Protein involved i	333	7	0.9	660	20	AAW13349	Amino acid sequenc
261	7	0.9	415	21	AAW37412	Arabidopsis thalia	334	7	0.9	660	22	AAU12346	Human PRO265 polyp
262	7	0.9	430	21	AAW26322	Arabidopsis thalia	335	7	0.9	660	22	AAW31208	Amino acid sequenc
263	7	0.9	430	21	AAW48079	Arabidopsis thalia	336	7	0.9	660	22	AAW80217	Human PRO265 prote
264	7	0.9	431	21	AAW37411	Arabidopsis thalia	337	7	0.9	661	18	AAW28510	Product of clone J
265	7	0.9	440	21	AAW22002	Arabidopsis thalia	338	7	0.9	661	19	AAW47274	Human B-cell activ
266	7	0.9	440	21	AAW46542	Arabidopsis thalia	339	7	0.9	661	20	AAW87556	B cell surface pro
267	7	0.9	447	21	AAW12574	NDO related comple	340	7	0.9	662	21	AAW14226	Arabidopsis thalia
268	7	0.9	447	21	AAW12575	NDO related comple	341	7	0.9	677	19	AAW55984	Swinepox virus str
269	7	0.9	451	21	AAW12576	NDO related comple	342	7	0.9	677	22	AAW59861	Drosophila melanog
270	7	0.9	454	22	ABB70036	Drosophila melanog	343	7	0.9	677	22	ABB61500	Drosophila melanog
271	7	0.9	455	22	AAE05769	Human secreted pro	344	7	0.9	677	22	ABB66398	Drosophila melanog
272	7	0.9	456	22	ABB61257	Drosophila melanog	345	7	0.9	697	21	AAW25565	Eucalyptus grandis
273	7	0.9	459	21	AAW20859	Arabidopsis thalia	346	7	0.9	705	21	AAW38743	Arabidopsis thalia
274	7	0.9	465	21	AAW14228	Drosophila melanog	347	7	0.9	708	21	AAW70473	Human cyclic nucle
275	7	0.9	474	22	AAW94199	Human protein sequ	348	7	0.9	711	21	AAW25111	Pinus radiata cell
276	7	0.9	474	22	AAE09322	Rice zinc transpo	349	7	0.9	711	20	AAW38742	Arabidopsis thalia
277	7	0.9	477	21	AAW25534	Arabidopsis thalia	350	7	0.9	713	20	AAW13385	Amino acid sequenc
278	7	0.9	480	21	AAW25533	Arabidopsis thalia	351	7	0.9	713	22	AAW80253	Human PRO293 prote
279	7	0.9	483	21	AAW91433	Human secreted pro	352	7	0.9	716	21	AAW76005	Rat Slit homologue
280	7	0.9	489	22	AAW92515	Human protein sequ	353	7	0.9	716	22	AAW55944	Skin cell protein,
281	7	0.9	492	22	ABB63749	Drosophila melanog	354	7	0.9	720	22	AAW75413	Human colon cancer
282	7	0.9	499	22	ABG25736	Novel human diagno	355	7	0.9	721	22	AAW93467	Human protein sequ
283	7	0.9	505	22	ABB58292	Drosophila melanog	356	7	0.9	723	22	AAU20426	Human secreted pro
284	7	0.9	512	16	AAW72364	Human auxillary cy	357	7	0.9	732	13	AAW20001	Human acyl amino a
285	7	0.9	512	16	AAW72365	Human auxillary cy	358	7	0.9	732	22	ABB44559	Human wound healin
286	7	0.9	512	16	AAW72366	Human auxillary cy	359	7	0.9	733	22	AAW70721	S cerevisiae apopt
287	7	0.9	512	17	AAW93172	Human cytochrome P	360	7	0.9	743	22	AAW00822	Human bone marrow
288	7	0.9	512	17	AAW93173	Human cytochrome P	361	7	0.9	753	22	AAW95109	Human protein sequ
289	7	0.9	512	17	AAW93171	Human cytochrome P	362	7	0.9	753	22	AAW95726	Human protein sequ
290	7	0.9	512	17	AAW00652	Cytochrome P450A1	363	7	0.9	768	22	ABB68660	Drosophila melanog
291	7	0.9	512	17	AAW93216	Human cytochrome P	364	7	0.9	783	10	AAW94265	Sequence of APH36.
292	7	0.9	512	22	AAW60566	Drosophila melanog	365	7	0.9	793	22	ABB58866	Drosophila melanog
293	7	0.9	521	21	AAW14227	Arabidopsis thalia	366	7	0.9	799	22	AAW78623	Human tyrosine pho
294	7	0.9	522	21	AAW43111	Human ORFX ORF2875	367	7	0.9	804	21	AAW25515	Pinus radiata cell
295	7	0.9	522	22	AAW78855	Human protein SEQ	368	7	0.9	807	20	AAW89247	Human PTP04, Homo
296	7	0.9	522	22	AAW25639	Human protein sequ	369	7	0.9	808	20	AAW28652	Human Cytoplasmic
297	7	0.9	526	22	AAW99895	Xanthomonas sp. ch	370	7	0.9	809	19	AAW21508	Human neuroendocri
298	7	0.9	543	22	AAW39906	Human polypeptide	371	7	0.9	811	20	AAW05868	Human Toll protein
299	7	0.9	544	20	AAW35764	Arabidopsis thalia	372	7	0.9	811	21	AAW66644	Membrane-bound pro
300	7	0.9	546	20	AAW13386	Amino acid sequenc	373	7	0.9	811	22	AAU29230	Human PRO polypept
301	7	0.9	546	21	AAW35763	Arabidopsis thalia	374	7	0.9	812	22	AAW65167	Human PRO358 prote
302	7	0.9	546	22	AAW25229	Human PRO polypept	375	7	0.9	828	21	AAW25559	Drosophila melanog
303	7	0.9	546	22	AAW93765	Human polypeptide,	376	7	0.9	843	21	AAW25518	Pinus radiata cell



523	6	0.7	24	12	AAR15026	Part of elastomeri	596	6	0.7	44	22	AAM87426	Human immune/haema
524	6	0.7	24	12	AAR15033	Part of elastomeri	597	6	0.7	44	22	AAM24523	Clone 27540 predic
525	6	0.7	24	16	AAR71299	GPV tandem Leu-ric	598	6	0.7	45	22	AAM85967	Human immune/haema
526	6	0.7	24	21	AAV56829	Apple LRPkml LRR s	599	6	0.7	45	22	RAO07102	Human polypeptide
527	6	0.7	24	21	AAV56846	Apple LRPkml LRR s	600	6	0.7	47	20	AAV02945	Fragment of human
528	6	0.7	25	18	AAW35364	Human 65K-glutamic	601	6	0.7	47	22	ABBA4279	Peptide #11785 enco
529	6	0.7	25	18	AAW30111	Hepatitis B e-anti	602	6	0.7	47	22	ABBA27144	Protein #9143 enco
530	6	0.7	25	18	AAW30112	Hepatitis B e-anti	603	6	0.7	47	22	AAM65313	Human brain expres
531	6	0.7	25	18	AAW30113	Hepatitis B e-anti	604	6	0.7	47	22	AAM78013	Human bone marrow
532	6	0.7	25	21	AAV64756	Human 5' EST relat	605	6	0.7	47	22	AAM21902	Peptide #8336 enco
533	6	0.7	26	12	AAR15012	Part of elastomeri	606	6	0.7	47	22	AAM38229	Peptide #12266 enco
534	6	0.7	26	12	AAR14998	Part of elastomeri	607	6	0.7	48	20	AAV11468	Human 5' EST secre
535	6	0.7	26	12	AAR14956	Part of elastomeri	608	6	0.7	48	21	AAV13298	Caenorhabditis ele
536	6	0.7	26	12	AAR14963	Part of elastomeri	609	6	0.7	48	22	ABB71673	Drosophila melanog
537	6	0.7	26	12	AAR14970	Part of elastomeri	610	6	0.7	48	22	ABG18561	Novel human diagno
538	6	0.7	26	12	AAR15005	Part of elastomeri	611	6	0.7	48	22	AAO10458	Human polypeptide
539	6	0.7	26	12	AAR14977	Part of elastomeri	612	6	0.7	49	21	AAV56775	Arabisopsis thalia
540	6	0.7	26	12	AAR15019	Part of elastomeri	613	6	0.7	49	22	ABB37392	Peptide #4898 enco
541	6	0.7	27	21	AAV34093	Human secreted pro	614	6	0.7	49	22	ABB38087	Peptide #5593 enco
542	6	0.7	27	22	ABBA41097	Peptide #8603 enco	615	6	0.7	49	22	ABBA1631	Peptide #9137 enco
543	6	0.7	27	22	ABBA25150	Protein #7149 enco	616	6	0.7	49	22	ABBA23684	Protein #4683 enco
544	6	0.7	27	22	ABBA11917	Human secreted pro	617	6	0.7	49	22	ABBA23308	Protein #5307 enco
545	6	0.7	27	22	AAM61953	Human brain expres	618	6	0.7	49	22	ABBA25445	Protein #7444 enco
546	6	0.7	27	22	AAW74757	Human bone marrow	619	6	0.7	49	22	AAW58724	Human brain expres
547	6	0.7	27	22	AAW34873	Peptide #8910 enco	620	6	0.7	49	22	AAW62502	Human brain expres
548	6	0.7	28	22	ABG04800	Novel human diagno	621	6	0.7	49	22	AAW70521	Human bone marrow
549	6	0.7	30	22	ABBA41439	Peptide #8945 enco	622	6	0.7	49	22	AAW71231	Human bone marrow
550	6	0.7	30	22	ABBA25339	Protein #7338 enco	623	6	0.7	49	22	AAW75311	Human bone marrow
551	6	0.7	30	22	AAM62312	Human brain expres	624	6	0.7	49	22	AAW18938	Peptide #5372 enco
552	6	0.7	30	22	AAW75115	Human bone marrow	625	6	0.7	49	22	AAW20565	Peptide #6999 enco
553	6	0.7	30	22	AAM20503	Peptide #6937 enco	626	6	0.7	49	22	AAW31512	Peptide #5549 enco
554	6	0.7	30	22	AAM35231	Peptide #9268 enco	627	6	0.7	49	22	AAW35424	Peptide #9461 enco
555	6	0.7	31	18	AAW24753	Finger 2 domain of	628	6	0.7	50	20	AAW89327	Neisseria meningit
556	6	0.7	33	21	AAW28749	Human secreted pro	629	6	0.7	50	21	AAW15193	Arabisopsis thalia
557	6	0.7	34	21	AAW16928	Bacteriophage Dp-1	630	6	0.7	50	22	AAW85617	Human immune/haema
558	6	0.7	35	21	AAW09432	Arabisopsis thalia	631	6	0.7	50	22	AAO08512	Human polypeptide
559	6	0.7	35	21	AAW59192	Arabisopsis thalia	632	6	0.7	52	16	AAW79646	pmON5969 (20-70)hi
560	6	0.7	35	21	AAW02760	Mouse GDF-1 finger	633	6	0.7	52	20	AAW48338	Human prostate can
561	6	0.7	35	21	AAW92531	Finger 2 subdomain	634	6	0.7	52	21	AAW16816	Bacteriophage Dp-1
562	6	0.7	36	20	AAW12179	Human 5' EST secre	635	6	0.7	52	22	ABG06521	Novel human diagno
563	6	0.7	37	20	AAW12732	Human 5' EST secre	636	6	0.7	52	22	ABG06521	Novel human diagno
564	6	0.7	37	20	AAW85094	Polypeptide fragme	637	6	0.7	52	22	ABBA38222	Peptide #5728 enco
565	6	0.7	37	21	AAW09493	Murine GDF-1 finge	638	6	0.7	52	22	AAW82163	Human haematologic
566	6	0.7	37	22	ABBA1268	Human secreted pro	639	6	0.7	52	22	AAW71372	Human bone marrow
567	6	0.7	37	22	AAW11819	Human polypeptide	640	6	0.7	53	21	AAW38068	Fragment of human
568	6	0.7	39	20	AAW12779	Human 5' EST secre	641	6	0.7	53	21	AAW04359	Arabisopsis thalia
569	6	0.7	39	22	AAW04471	Human polypeptide	642	6	0.7	53	21	AAW07785	Arabisopsis thalia
570	6	0.7	40	22	AAW13103	Human secreted pro	643	6	0.7	54	20	AAW60377	Human normal bladd
571	6	0.7	40	22	ABBA31680	Peptide #4331 enco	644	6	0.7	54	21	AAW02619	Human secreted pro
572	6	0.7	40	22	ABBA36895	Peptide #4401 enco	645	6	0.7	54	22	AAW90397	Human immune/haema
573	6	0.7	40	22	ABBA22218	Protein #4217 enco	646	6	0.7	54	22	AAW76889	Human colon cancer
574	6	0.7	40	22	AAW57640	Human brain expres	647	6	0.7	55	20	AAW74192	Human prostate tum
575	6	0.7	40	22	AAW70045	Human bone marrow	648	6	0.7	55	20	AAV02902	Fragment of human
576	6	0.7	40	22	AAW04903	Human polypeptide	649	6	0.7	55	21	AAW56774	Arabisopsis thalia
577	6	0.7	40	22	AAW09095	Human polypeptide	650	6	0.7	55	22	ABG08787	Novel human diagno
578	6	0.7	40	22	AAW17876	Peptide #4310 enco	651	6	0.7	55	22	AAW84497	Human immune/haema
579	6	0.7	40	22	AAW30383	Peptide #4420 enco	652	6	0.7	55	22	AAW85309	Human immune/haema
580	6	0.7	40	22	AAW05524	Peptide #4206 enco	653	6	0.7	56	21	AAW02413	Human secreted pro
581	6	0.7	41	21	AAV76206	Human secreted pro	654	6	0.7	56	21	AAV32152	Human secreted pro
582	6	0.7	41	22	AAW87107	Human immune/haema	655	6	0.7	56	22	AAW47542	Propionibacterium
583	6	0.7	41	22	AAW66733	Aldehyde/ketone re	656	6	0.7	56	22	ABBA16917	Human nervous syst
584	6	0.7	42	14	AAW39731	C-terminal type I	657	6	0.7	56	22	AAW78296	Human bone marrow
585	6	0.7	42	18	AAW19979	Fibronectin fragme	658	6	0.7	56	22	AAW82602	Human immune/haema
586	6	0.7	42	20	AAW83001	Human fibronectin	659	6	0.7	57	22	AAU66084	Propionibacterium
587	6	0.7	42	22	ABBA2284	Peptide #9790 enco	660	6	0.7	57	22	ABBA35806	Peptide #3312 enco
588	6	0.7	42	22	AAW63170	Human brain expres	661	6	0.7	57	22	ABBA16664	Human nervous syst
589	6	0.7	42	22	AAW75984	Human bone marrow	662	6	0.7	57	22	AAW68991	Human bone marrow
590	6	0.7	42	22	AAW36092	Peptide #10129 enco	663	6	0.7	57	22	AAW83824	Human immune/haema
591	6	0.7	42	22	AAE06107	Human gene 67 enco	664	6	0.7	57	22	AAW16822	Peptide #3256 enco
592	6	0.7	43	20	AAV14019	YhdD repeat sequen	665	6	0.7	57	22	AAW29305	Peptide #3342 enco
593	6	0.7	43	21	AAW87130	Human secreted pro	666	6	0.7	57	22	AAW04534	Peptide #3216 enco
594	6	0.7	44	22	ABG28609	Novel human diagno	667	6	0.7	58	16	AAW80870	z protein variant
595	6	0.7	44	22	AAU21450	Human novel foetal	668	6	0.7	58	22	AAW99727	ERA binding domain

669	6	0.7	59	20	AA36177	Human secreted pro	742	6	0.7	72	22	ABB52596	Escherichia coli p
670	6	0.7	59	21	AA61003	Arabidopsis thalia	743	6	0.7	72	22	AAU32986	Novel human secret
671	6	0.7	59	22	AA63717	Propionibacterium	744	6	0.7	72	22	AAO04100	Human polypeptide
672	6	0.7	59	22	AA95951	Human reproductive	745	6	0.7	73	21	AA44406	Gene 43 encoded hu
673	6	0.7	60	21	AA8076	Fragment of human	746	6	0.7	73	22	AAU56149	Propionibacterium
674	6	0.7	60	22	AAU17628	Novel signal trans	747	6	0.7	73	22	AAU60426	Propionibacterium
675	6	0.7	61	21	AA611142	Arabidopsis thalia	748	6	0.7	73	22	ABG14062	Novel human diago
676	6	0.7	61	22	AA84372	Human immune/haema	749	6	0.7	73	22	ABB16595	Human nervous syst
677	6	0.7	61	22	AAU14679	Novel bone marrow	750	6	0.7	73	22	AA886445	Human immune/haema
678	6	0.7	62	22	AA899040	Goldfish GnRH rece	751	6	0.7	73	22	AAW91734	Human immune/haema
679	6	0.7	62	18	AAW20505	H. pylori cytoplas	752	6	0.7	73	22	AAO08124	Human polypeptide
680	6	0.7	62	22	AAU55209	Propionibacterium	753	6	0.7	73	22	AAE03078	Human gene 1 encod
681	6	0.7	62	22	AAU27335	Novel bone marrow	754	6	0.7	74	21	AAAG01846	Human secreted pro
682	6	0.7	63	20	AA28478	Partial amino acid	755	6	0.7	74	22	AAU66109	Human secreted pro
683	6	0.7	63	20	AA11452	Human 5' EST sece	756	6	0.7	75	15	AAU56155	Human interleukin-
684	6	0.7	63	22	AB44131	Peptide #11637 enc	757	6	0.7	75	16	AA69741	pMON5917 (Met-15-8
685	6	0.7	63	22	AB27013	Protein #9012 enco	758	6	0.7	75	19	AAW74920	Human secreted pro
686	6	0.7	63	22	AA65154	Human brain expres	759	6	0.7	75	22	ABG21489	Novel human diago
687	6	0.7	63	22	AAW77859	Human bone marrow	760	6	0.7	75	22	AAAG0617	Human immune/haema
688	6	0.7	63	22	AAW21759	Peptide #8193 enco	761	6	0.7	76	20	AA15212	Sphingosine-1-phos
689	6	0.7	63	22	AA38081	Peptide #12118 enc	762	6	0.7	76	22	ABB71679	Drosophila melanog
690	6	0.7	64	16	AA68557	Klebsiella sp. nit	763	6	0.7	76	22	ABG16073	Novel human diago
691	6	0.7	65	22	AAU58542	Propionibacterium	764	6	0.7	76	22	ABG12172	Human MECHP-5 homo
692	6	0.7	65	22	ABG23564	Novel human diago	765	6	0.7	77	20	AAU06659	Magneporth grisea
693	6	0.7	65	22	AB43360	Peptide #10866 enc	766	6	0.7	77	20	AAU02901	Fragment of human
694	6	0.7	65	22	AB17350	Human nervous syst	767	6	0.7	77	22	AAU65306	Propionibacterium
695	6	0.7	65	22	AAW77094	Human bone marrow	768	6	0.7	77	22	ABG30209	Novel human diago
696	6	0.7	66	22	AAU45773	Propionibacterium	769	6	0.7	77	22	ABB23301	Peptide #1932 enco
697	6	0.7	66	22	AAU51503	Propionibacterium	770	6	0.7	77	22	ABB34473	Peptide #1979 enco
698	6	0.7	66	22	AA86050	Human immune/haema	771	6	0.7	77	22	AB440530	Peptide #8036 enco
699	6	0.7	66	22	AAU17617	Novel signal trans	772	6	0.7	77	22	AB119883	Protein #1882 enco
700	6	0.7	67	22	AB68172	Drosophila melanog	773	6	0.7	77	22	AB24850	Protein #6849 enco
701	6	0.7	67	22	AAU59583	Propionibacterium	774	6	0.7	77	22	AAU55259	Human brain expres
702	6	0.7	67	22	AB03420	Human musculoskele	775	6	0.7	77	22	AA61356	Human brain expres
703	6	0.7	67	22	AA885679	Salmonella virulen	776	6	0.7	77	22	AAW67656	Human bone marrow
704	6	0.7	67	22	AAO12284	Human polypeptide	777	6	0.7	77	22	AAW74094	Human bone marrow
705	6	0.7	68	22	AAU43529	Propionibacterium	778	6	0.7	77	22	AAO03897	Human polypeptide
706	6	0.7	68	22	ABG16517	Novel human diago	779	6	0.7	77	22	AAU15464	Peptide #1898 enco
707	6	0.7	68	22	AA80077	Human protein SEQ	780	6	0.7	77	22	AAU27952	Peptide #1989 enco
708	6	0.7	69	20	AAU59645	Secreted protein e	781	6	0.7	77	22	AAU34256	Peptide #8293 enco
709	6	0.7	69	20	AAU36172	Human secreted pro	782	6	0.7	77	22	AAU03221	Peptide #1903 enco
710	6	0.7	69	20	AAU36219	Human secreted pro	783	6	0.7	78	20	AAU48319	Human prostate can
711	6	0.7	69	20	AAU25451	Human secreted pro	784	6	0.7	78	22	AAU41173	Propionibacterium
712	6	0.7	69	20	AAU35882	Extended human sec	785	6	0.7	78	22	ABG16074	Novel human diago
713	6	0.7	69	20	AAW93616	Human 5' EST sece	786	6	0.7	78	22	AB37806	Peptide #5312 enco
714	6	0.7	69	20	AAU12672	Human 5' EST sece	787	6	0.7	78	22	AB23081	Peptide #5080 enco
715	6	0.7	69	20	AAU12982	Human secreted pro	788	6	0.7	78	22	AAU58422	Human brain expres
716	6	0.7	69	20	AAU04166	Human 5' EST sece	789	6	0.7	78	22	AAW70911	Human bone marrow
717	6	0.7	69	20	AAU01590	Secreted protein e	790	6	0.7	78	22	AAW87738	Human immune/haema
718	6	0.7	69	20	AAU04148	Human 5' EST sece	791	6	0.7	78	22	AAO03565	Human polypeptide
719	6	0.7	69	21	AA42587	Human ORFX ORF2351	792	6	0.7	78	22	AAO07977	Human polypeptide
720	6	0.7	69	21	AAO05284	Arabidopsis thalia	793	6	0.7	78	22	AAU31203	Peptide #5240 enco
721	6	0.7	69	21	AAU00010	Human secreted pro	794	6	0.7	78	22	AAU74141	Human colon cancer
722	6	0.7	69	21	AAU02835	Human secreted pro	795	6	0.7	79	21	AAU15170	Arabidopsis thalia
723	6	0.7	69	22	AAU63290	Propionibacterium	796	6	0.7	79	21	AAU03231	Human secreted pro
724	6	0.7	70	20	AAU24002	A flea epoxide hyd	797	6	0.7	79	22	ABG24153	Novel human diago
725	6	0.7	70	20	AAU24006	A flea epoxide hyd	798	6	0.7	79	22	AAU83178	Human immune/haema
726	6	0.7	70	21	AAU20926	Arabidopsis thalia	799	6	0.7	79	22	AAO02285	Human polypeptide
727	6	0.7	70	21	AAU21321	Arabidopsis thalia	800	6	0.7	79	22	AAU77339	Human colon cancer
728	6	0.7	70	21	AAU58849	Arabidopsis thalia	801	6	0.7	79	22	AAU43831	Propionibacterium
729	6	0.7	70	22	AAU65734	Flea epoxide hydro	802	6	0.7	80	22	ABG11056	Novel human diago
730	6	0.7	70	22	AAU65738	Flea epoxide hydro	803	6	0.7	80	22	AAW27848	Novel human diago
731	6	0.7	70	22	AAU51004	Flea epoxide hydro	804	6	0.7	81	18	AAW75962	Human Caspase-81.
732	6	0.7	70	22	AAU51008	Flea epoxide hydro	805	6	0.7	81	21	AAU01309	Human secreted pro
733	6	0.7	70	22	AAU36981	Flea nFEH1 211 pro	806	6	0.7	81	21	AAU01311	Human secreted pro
734	6	0.7	70	22	AAU36985	nFEH2 211 protein.	807	6	0.7	82	21	AAU020420	Arabidopsis thalia
735	6	0.7	71	22	AAU15266	Human nervous syst	808	6	0.7	82	22	AAE03116	Human gene 20 enco
736	6	0.7	71	22	AAU15968	Human nervous syst	809	6	0.7	82	22	AAU41787	Propionibacterium
737	6	0.7	71	22	AAU18203	Novel human DNA-bi	810	6	0.7	83	22	AAU54745	Propionibacterium
738	6	0.7	71	22	AAU90899	Human immune/haema	811	6	0.7	83	22	AAU55551	Propionibacterium
739	6	0.7	71	22	AAO08969	Human polypeptide	812	6	0.7	83	22	AAO04536	Human polypeptide
740	6	0.7	71	22	AAU13297	Human polypeptide	813	6	0.7	83	22	AAU74906	Human colon cancer
741	6	0.7	71	22	AAU74469	Human colon cancer	814	6	0.7	83	22	AAU84796	Caspase 8 death ef

815	6	0.7	83	22	AAB61111	Human caspase-8 DE	888	6	0.7	98	22	AAM54758	Human brain expres
816	6	0.7	84	21	AAG54966	Arabidopsis thalia	889	6	0.7	98	22	AAG74915	Human colon cancer
817	6	0.7	84	21	AG02066	Human secreted pro	890	6	0.7	99	19	AAW59429	Human chemokine pr
818	6	0.7	84	22	AAU41990	Propionibacterium	891	6	0.7	100	18	AAW28090	Amino acid sequenc
819	6	0.7	84	22	AAU53808	Propionibacterium	892	6	0.7	100	22	AAU56132	Propionibacterium
820	6	-0.7	84	22	AAU58467	Human brain expres	893	6	0.7	100	22	AAU63782	Novel human thalia
821	6	0.7	84	22	AAW70961	Human bone marrow	894	6	0.7	101	22	ABG03168	Arabidopsis thalia
822	6	0.7	85	21	AAU86268	Human secreted pro	895	6	0.7	101	22	AAU60269	Human secreted pro
823	6	0.7	85	22	ABG18929	Novel human diagno	896	6	0.7	101	22	AAU03822	Human immune/haema
824	6	0.7	85	22	ABG28360	Peptide #1011 enco	897	6	0.7	101	22	AAU85263	Human lung tumour
825	6	0.7	85	22	ABG33539	Peptide #1045 enco	898	6	0.7	102	20	AAU29549	Human lung tumour
826	6	0.7	85	22	ABG38730	Peptide #6236 enco	899	6	0.7	102	21	AAU44472	Zea mays protein f
827	6	0.7	85	22	ABG18995	Protein #994 enco	900	6	0.7	102	21	AAU32690	Human lung tumour
828	6	0.7	85	22	AAU31403	Novel human secret	901	6	0.7	102	22	AAU13814	Human polypeptide
829	6	0.7	85	22	AAU54315	Human brain expres	902	6	0.7	102	22	AAU12332	Human foetal prote
830	6	0.7	85	22	AAU59364	Human brain expres	903	6	0.7	102	22	AAU06407	Chlamydia trachoma
831	6	0.7	85	22	AAU66713	Human bone marrow	904	6	0.7	103	20	AAU37928	Arabidopsis thalia
832	6	0.7	85	22	AAU71919	Human bone marrow	905	6	0.7	103	21	AAU11141	Arabidopsis thalia
833	6	0.7	85	22	AAU09990	Human immune/haema	906	6	0.7	103	21	AAU46644	Arabidopsis thalia
834	6	0.7	85	22	AAU010195	Human polypeptide	907	6	0.7	103	22	AAU50662	Propionibacterium
835	6	0.7	85	22	AAU14578	Peptide #1012 enco	908	6	0.7	104	21	AAU40540	Arabidopsis thalia
836	6	0.7	85	22	AAU26999	Peptide #1036 enco	909	6	0.7	104	22	AAU83443	Human immune/haema
837	6	0.7	85	22	AAU32200	Peptide #6237 enco	910	6	0.7	105	16	AAU69761	PMON5974 (Met-15-1
838	6	0.7	85	22	AAU02304	Peptide #986 enco	911	6	0.7	105	21	AAU22875	Arabidopsis thalia
839	6	0.7	86	18	AAU35360	Mouse 65K-glutamic	912	6	0.7	105	21	AAU39724	Arabidopsis thalia
840	6	0.7	86	22	AAU95059	Human reproductive	913	6	0.7	105	22	AAU07500	Human polypeptide
841	6	0.7	87	21	AAU65228	Human 5' EST relat	914	6	0.7	106	18	AAU20984	H. pylori cytoplas
842	6	0.7	87	22	AAU89444	Human immune/haema	915	6	0.7	106	21	AAU75958	Rat skin cell prot
843	6	0.7	87	22	AAU04290	Human polypeptide	916	6	0.7	106	22	AAU55897	Skin cell protein,
844	6	0.7	87	22	AAU09569	Human polypeptide	917	6	0.7	107	17	AAU03173	STAR protein DNA b
845	6	0.7	89	21	AAU01464	Human secreted pro	918	6	0.7	107	21	AAU09537	Murine GDF-1, SEQ
846	6	0.7	89	22	AAU36458	Pseudomonas aerugi	919	6	0.7	107	21	AAU02803	Mouse GDF-1 amino
847	6	0.7	89	22	AAU84334	Human immune/haema	920	6	0.7	107	21	AAU92572	GDF 1 finger-1-hee
848	6	0.7	89	22	AAU01004	Human polypeptide	921	6	0.7	107	22	AAU79597	Human protein SEQ
849	6	0.7	90	21	AAU35351	Zea mays protein f	922	6	0.7	107	22	AAU06855	Human foetal prote
850	6	0.7	91	21	AAU28948	Wheat auxin transp	923	6	0.7	107	22	AAU33902	Arabidopsis thalia
851	6	0.7	91	22	AAU58147	Propionibacterium	924	6	0.7	108	22	AAU53811	Propionibacterium
852	6	0.7	91	22	ABG16071	Novel human diagno	925	6	0.7	108	22	AAU55867	Propionibacterium
853	6	0.7	91	22	AAU17165	Novel signal trans	926	6	0.7	109	21	AAU25366	Pinus radiata cell
854	6	0.7	92	22	AAU24114	Human EST encoded	927	6	0.7	109	22	ABG00973	Novel human diagno
855	6	0.7	92	22	AAU61804	Propionibacterium	928	6	0.7	109	22	AAU02329	Human polypeptide
856	6	0.7	92	22	AAU18958	Novel lung cancer	929	6	0.7	109	22	AAU11357	Human polypeptide
857	6	0.7	92	22	AAU17973	Novel human respir	930	6	0.7	109	22	AAU13230	Human polypeptide
858	6	0.7	93	18	AAU20334	H. pylori cytoplas	931	6	0.7	110	16	AAU69757	PMON5968 (Met-15-1
859	6	0.7	93	19	AAU44831	Oxytocin precursor	932	6	0.7	110	21	AAU34961	Gene 33 human secr
860	6	0.7	93	21	AAU34887	Arabidopsis thalia	933	6	0.7	110	22	ABG30659	Peptide #3310 enco
861	6	0.7	93	22	AAU51317	Propionibacterium	934	6	0.7	110	22	ABG35830	Peptide #3336 enco
862	6	0.7	93	22	ABG11137	Human zn finger pr	935	6	0.7	110	22	ABG21245	Protein #3244 enco
863	6	0.7	94	20	AAU12768	Human 5' EST, secre	936	6	0.7	110	22	AAU56636	Human brain expres
864	6	0.7	94	21	AAU332691	Zea mays protein f	937	6	0.7	110	22	AAU69013	Human bone marrow
865	6	0.7	94	21	AAU37160	Arabidopsis thalia	938	6	0.7	110	22	AAU84762	Human immune/haema
866	6	0.7	94	21	AAU01526	Human secreted pro	939	6	0.7	110	22	AAU85270	Human immune/haema
867	6	0.7	94	22	ABG15548	Human nervous syst	940	6	0.7	110	22	AAU15847	Peptide #3281 enco
868	6	0.7	94	22	AAU95715	Human reproductive	941	6	0.7	110	22	AAU29332	Peptide #3369 enco
869	6	0.7	95	20	AAU35929	Extended human sec	942	6	0.7	110	22	AAU04557	Peptide #3239 enco
870	6	0.7	95	21	AAU15169	Arabidopsis thalia	943	6	0.7	110	22	AAU91710	C glutamic prote
871	6	0.7	95	22	AAU49715	Propionibacterium	944	6	0.7	111	15	AAU56158	Human interleukin-
872	6	0.7	95	22	AAU95844	Human reproductive	945	6	0.7	111	15	AAU56012	Sequence of mutant
873	6	0.7	95	22	AAU31105	Peptide #5142 enco	946	6	0.7	111	15	AAU56020	Sequence of mutant
874	6	0.7	96	20	AAU33803	Chlamydia pneumoni	947	6	0.7	111	15	AAU56021	Sequence of mutant
875	6	0.7	96	21	AAU26951	Wheat auxin transp	948	6	0.7	111	15	AAU56047	Human interleukin-
876	6	0.7	96	21	AAU10447	Arabidopsis thalia	949	6	0.7	111	15	AAU56048	Human interleukin-
877	6	0.7	96	21	AAU58824	Arabidopsis thalia	950	6	0.7	111	15	AAU56049	Human interleukin-
878	6	0.7	96	22	AAU20858	Human novel foetal	951	6	0.7	111	15	AAU56050	Human interleukin-
879	6	0.7	96	22	AAU00428	Human polypeptide	952	6	0.7	111	15	AAU56053	Human interleukin-
880	6	0.7	97	19	AAU60929	Streptococcus pneu	953	6	0.7	111	15	AAU56056	Human interleukin-
881	6	0.7	97	22	AAU08492	Mouse VMGIOM short	954	6	0.7	111	15	AAU56054	Human interleukin-
882	6	0.7	98	21	AAU11346	Arabidopsis thalia	955	6	0.7	111	15	AAU56051	Human interleukin-
883	6	0.7	98	21	AAU44435	Arabidopsis thalia	956	6	0.7	111	15	AAU56055	Human interleukin-
884	6	0.7	98	21	AAU51292	Arabidopsis thalia	957	6	0.7	111	15	AAU56052	Human interleukin-
885	6	0.7	98	21	AAU57723	Arabidopsis thalia	958	6	0.7	111	16	AAU79257	Mutant human inter
886	6	0.7	98	22	ABG28808	Peptide #1459 enco	959	6	0.7	111	16	AAU79258	Mutant human inter
887	6	0.7	98	22	ABG19433	Protein #1432 enco	960	6	0.7	111	16	AAU79259	Mutant human inter

961 6 0.7 111 16 AAR79260 Mutant human inter  
962 6 0.7 111 16 AAR79261 Mutant human inter  
963 6 0.7 111 16 AAR79262 Mutant human inter  
964 6 0.7 111 16 AAR79263 Mutant human inter  
965 6 0.7 111 16 AAR79264 Mutant human inter  
966 6 0.7 111 16 AAR79265 Mutant human inter  
967 6 0.7 111 16 AAR79266 Mutant human inter  
968 6 0.7 111 16 AAR79332 Human IL-3 mutant  
969 6 0.7 111 16 AAR79333 Human IL-3 mutant  
970 6 0.7 111 16 AAR79334 Human IL-3 mutant  
971 6 0.7 111 16 AAR79335 Human IL-3 mutant  
972 6 0.7 111 16 AAR79336 Human IL-3 mutant  
973 6 0.7 111 16 AAR79337 Human IL-3 mutant  
974 6 0.7 111 16 AAR79338 Human IL-3 mutant  
975 6 0.7 111 16 AAR79339 Human IL-3 mutant  
976 6 0.7 111 16 AAR79340 Human IL-3 mutant  
977 6 0.7 111 16 AAR79341 Human IL-3 mutant  
978 6 0.7 111 21 AAY53141 Human interleukin-  
979 6 0.7 111 21 AAY53142 Human interleukin-  
980 6 0.7 111 21 AAY53143 Human interleukin-  
981 6 0.7 111 21 AAY53144 Human interleukin-  
982 6 0.7 111 21 AAY53145 Human interleukin-  
983 6 0.7 111 21 AAY53146 Human interleukin-  
984 6 0.7 111 21 AAY53147 Human interleukin-  
985 6 0.7 111 21 AAY53148 Human interleukin-  
986 6 0.7 111 21 AAY53149 Human interleukin-  
987 6 0.7 111 21 AAY53150 Human interleukin-  
988 6 0.7 111 22 AAE13863 Myelopoietin (MPO)  
989 6 0.7 111 22 AAE13864 Myelopoietin (MPO)  
990 6 0.7 111 22 AAE13865 Myelopoietin (MPO)  
991 6 0.7 111 22 AAE13866 Myelopoietin (MPO)  
992 6 0.7 111 22 AAE13867 Myelopoietin (MPO)  
993 6 0.7 111 22 AAE13868 Myelopoietin (MPO)  
994 6 0.7 111 22 AAE13869 Myelopoietin (MPO)  
995 6 0.7 111 22 AAE13870 Myelopoietin (MPO)  
996 6 0.7 111 22 AAE13871 Myelopoietin (MPO)  
997 6 0.7 111 22 AAE13872 Myelopoietin (MPO)  
998 6 0.7 111 22 AAM23832 Human EST encoded  
999 6 0.7 112 15 AAR56157 Human interleukin-  
1000 6 0.7 112 15 AAR56154 Human interleukin-

ALIGNMENTS

RESULT 1  
AAW86354  
ID AAW86354 standard; Protein; 1045 AA.

XX AAW86354;  
XX  
XX  
XX 15-MAR-1999 (first entry)  
XX  
XX

DE Human DNAX toll-like receptor DTLR6.

XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;  
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;  
KW modulate inflammatory function; morphological effect;  
KW immunological disorder.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH Peptide 1..22  
FT /label= signal  
FT Protein 23..1045  
FT /label= DTLR6

XX W09850547-A2.

XX 12-NOV-1998.

XX 07-MAY-1998; 98WO-US08979.

XX 05-MAR-1998; 98US-0076947.  
PR 07-MAY-1997; 97US-0044293.  
PR 22-JAN-1998; 98US-0072212.  
XX (SCHE ) SCHERING CORP.  
PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;  
XX WPI; 1999-059670/05.  
DR N-PSDB; AAV80868.  
XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter  
PT phosphate metabolism, modulate inflammatory function or innate  
PT immunity responses  
XX Claim 5; Page 125-128; 171pp; English.  
PS The present invention specifically describes human DNAX toll-like  
XX receptors 2 to 10 (DTLR2-10). The present sequence is human DTLR6  
CC given in the present invention. Also described are: (1) a fusion protein  
CC comprising a DTLR protein or peptide; (2) a binding compound, preferably  
CC an antibody or antibody fragment which specifically binds to a DTLR  
CC protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
CC peptide; (4) an expression vector comprising the nucleic acid of (3);  
CC and (5) a host cell comprising the vector of (4). The host cell of (5)  
CC can be used to produce the DTLR proteins. The DTLR proteins can be used  
CC to alter phosphate metabolism, to modulate inflammatory function, innate  
CC immunity responses or morphological effects. The DTLR proteins can be  
CC used in the treatment of conditions exhibiting abnormal expression of  
CC the receptors of their ligands. These abnormalities are typically  
CC manifested by immunological disorders.  
XX Sequence 1045 AA;  
SQ

Query Match 100.0%; Score 807; DB 20; Length 1045;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPKTLPDVTLDVPRKHHVTDCTDKHLTEIPGGIPTNTNLTINHIIDISPASFRHLD 60  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
26 fpktpcdvtdldvpknhvivdctdkhlteipggiptntnltinhpidspesfhrid 85  
QY 61 HLVEIDFRNCVPIPLGSKNNMCIKRLQIKPRFSGLTLYLKSLYLDGNOLLETPQGLPPS 120  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
86 hlveidfrncvpiplgsknnmcikrqlkprfsgltlylkslyldgnllepqlppps 145  
QY 121 LQLLSLEANNIFSRKENITELANIEILYLGQNCYYRNPICYYSYSTERDAFLNLTKLKVL 180  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
146 lqlslleannifsrkenitelanietellylgqncyyrnpicyvysylekdaflnltklvl 205  
QY 181 SLKDNNTAVPTVPLSTLTLYLYNNMIAKIQEDDFNNLNQIQLDLSGNCPCYNAPPP 240  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
206 slkdnnvtavptvplstltelylynnmiakiqeddfnnlnqilqlsgncpcynapfp 265  
QY 241 CAPCKNNSPLQIPVNAFDALTELKVLRLHNSLQHVPPRWFKNINKLQELDLSONFLAKE 300  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
266 capcknnsplqipvnafdaltelkvlrlhnsnlqhvpprwfkninklqeldlsonflake 325  
QY 301 IGDAKFLHFLPSLTQLDLSFNFELOVYRASMLNSQAFSSLSKLILIRIGYVFKELKSPN 360  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
326 igdakflhflpsliqlidsfnfelqvyrasmnlseqafsslsklilrirygvfkelksfn 385  
QY 361 LSPLHNLQNLVLDLGTNFIKIANLSMFQKRLKVIDLSVKNKISPSGDSSEVGFCSNAR 420  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
386 lsplhnlqlnlevidlgtntfikianismfkgfkrkvidlsvknkispdsgdssevgfcsnar 445  
QY 421 TSVESYEPQVLEQLHYFRYDKYARSCRFRNKKEASPMVNESCYKYGQTLDLKSNSTFFVK 480  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
446 tsvesyepqvlleqlhyfrydkyarscrfrnkkeaspmvnesckykygtldlksnstffvk 505  
QY 481 SSDFOHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRLDLLHSTAFEELHKLKE 540





PS Claim 12; Fig 209; 530pp; English.

XX The present invention describes secreted and transmembrane polypeptides  
CC and their polynucleotides. The nucleotide sequences are useful as  
CC sources of probes, primers, for chromosome mapping, and for generation  
CC of antisense sequences. They can also be used to create transgenic  
CC animals. The proteins can be used to treat a variety of diseases and  
CC disorders, depending on their function. Diseases that may be treated  
CC include blood coagulation disorders, cancers and cellular adhesion  
CC disorders. They may also be used to raise antibodies. A233891 to  
CC A234338, and AAY41685 to AAY41774 represent polynucleotide and  
CC polypeptide sequence given in the exemplification of the present  
XX invention.

SQ Sequence 1049 AA;

Query Match 100.0%; Score 807; DB 20; Length 1049;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPKTLPCDVTLDVKNHVVCTDKKHLTEIPGGIPTNTLTNTLNHPTDISPASFRHLD 60  
DB |||||||  
QY 30 fpktpcdvtdvknhvvctdkkhlteipggiptntltntlnhptdispasfrhld 89  
DB |||||||  
QY 61 HLVEIDFRNCVPIPLGSKNNMKIKRLQIKPRFSGLTYLKSLYLDGNQLLEIPQGLPPS 120  
DB |||||||  
QY 90 hliveidfrncvpipplgsknnmkirkqprsfsgltykslyldgnqlleipqglpps 149  
DB |||||||  
QY 121 LQLLSLEANNFSIRKENLTLANTEIILYLGNCYRNPVYSYIEKDANLTKLKV 180  
DB |||||||  
QY 150 lqlslsleannfisirkenlteianleilylgncyrynpvysyiekdafnltklklv 209  
DB |||||||  
QY 181 SLKDNNTAVPTLSTELYNNMKTAKIOEDDFNNLNLOIILDSGNCPCRYNAPFP 240  
DB |||||||  
QY 210 slkdnntavpctlstelynnmlakioeddfnnlnloilldsgncpcrynapfp 269  
DB |||||||  
QY 241 CAPKNNSPLOTPVNAFDALTELKVLRLHNSLSQHPVPRFKNINKLOELDLSQNLAKE 300  
DB |||||||  
QY 270 capknnsplotpvnafdalteklvrlhnsnlsqhpvrfrkninkloeldlsqnlake 329  
DB |||||||  
QY 301 GDAPKFLHFLSLIQLDLSFNFELQVYRASNNLSQAFSSLSKLTILRTGRVYFKELKSFN 360  
DB |||||||  
QY 330 igdakflhflpsliqlidsfnfelqvyrasmnlsqafsslskltkilrgyvfkelksfn 389  
DB |||||||  
QY 361 LSPHLNLQNLVLDLGTNFIKIANLSMFKQKRLKVIDLSYNNKISPSGDSSEVGFCSNAR 420  
DB |||||||  
QY 390 lsphlnlqnlvldlgtnfikianlsmfkqkrlkvidlsyNNKISPSGDSSEVGFCSNAR 449  
DB |||||||  
QY 421 TSVESYEPQVLEQLHYFRYDYKARSCRFKNKEASPMVSNECYKYQOTLDLSKNSIFPVK 480  
DB |||||||  
QY 450 tsvesyepqvleqlhyfrydykarscrfknkeasfmsvnescykygqtdlsknsifvfk 509  
DB |||||||  
QY 481 SSDFOHLSFLKCLNLUGNLISQTLNGSFPQIAELRYLDNNSNNRDLHSAFELHKLKLE 540  
DB |||||||  
QY 510 ssdfqhlslfklcnlsnqlisqtlngsefqplaelryldfnnrldllhstafeelhlkle 569  
DB |||||||  
QY 541 VLDISSNSHYFQSEGITHMLNFTKNLYLOKLMWMDNDISSSTSTMESESIRTLFPRGN 600  
DB |||||||  
QY 570 vldissnsHYFQSEGITHMLNFTKNLYLOKLMWMDNDISSSTSTMESESIRTLFPRGN 629  
DB |||||||  
QY 601 HLDVLWREGDNRYLQLFKNLKLEELDLSKNSLSPLSGVDFGMPNKNLSLAKNGLKS 660  
DB |||||||  
QY 630 hldvlwregdnrylqlfknllkeeldlsknslsplsgvdfgmpnknlslnknlgks 689  
DB |||||||  
QY 661 FSWKKLQCLKNLETLDLSHNLQTTVPRLSNCSRLKMLILKNNOIRSLTKYFLQDAFOL 720  
DB |||||||  
QY 690 fswkkqlcklnletldlshnqlttvprlsncsrslkmlilknnoirsaltkyflqdafql 749  
DB |||||||  
QY 721 RYLDLSSNKKIQMIOKTSFPENVLNNKMLLLHHNPFELCTCAVWFVWVWVNTETIPIYA 780  
DB |||||||  
QY 750 ryldlssnkkIQMIOKTSFPENVLNNKMLLLHHNPFELCTCAVWFVWVWVNTETIPIYA 809

Qy 781 TDVTCVGPCHAKGQSVISLDLYTCELD 807  
Db |||||||  
810 tdvtcvpggahkgqsvisdlytcelld 836  
RESULT 3  
AAY05866  
ID AAY05866 standard; Protein; 1049 AA.  
XX  
AC AAY05866;  
XX  
DT 02-AUG-1999 (first entry)  
XX  
DE Human Toll protein PRO285.  
XX  
KW PRO285; Toll; homologue; human; adaptive immunity; septic shock;  
KW inflammation; diabetes; amyotrophic lateral sclerosis; cancer;  
KW ulcer; rheumatoid arthritis; pathogen pattern recognition receptor;  
KW signal transduction.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..29  
FT Protein /note= "signal peptide"  
FT Protein 30..1049  
FT Domain /note= "mature protein"  
FT Domain 837..860  
FT Region /note= "transmembrane domain"  
FT Region 132..153  
FT Region /note= "leucine zipper"  
FT Region 704..725  
FT Modified-site /note= "leucine zipper"  
FT Modified-site 66  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 69  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 167  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 202  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 215  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 361  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 413  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 488  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 523  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 534  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 590  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 679  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 720  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 799  
FT Modified-site /note= "N-glycosylated"  
FT Modified-site 942  
FT Modified-site /note= "N-glycosylated"  
XX  
FN WO9920756-A2.  
XX  
PD 29-APR-1999.  
XX  
PF 07-OCT-1998; 98WO-US21141.  
XX  
PR 26-JUN-1998; 98US-0105413.  
PR 17-OCT-1997; 97US-0062250.  
PR 13-NOV-1997; 97US-0065311.

PR 28-APR-1998; 98US-0083322.  
PR 26-JUN-1998; 98US-0090863.  
PA (GETH ) GENENTECH INC.  
PI Goddard A, Godowski PJ, Gurney AL, Mark MR, Yang R;  
XX WPI; 1999-302739/25.  
DR N-PSDB; AAX58295.  
XX New human Toll-like receptors that recognize microbial structures  
PT Claim 1: Fig 1; 79pp; English.  
XX This is the amino acid sequence of PRO285, a novel human homologue  
CC of Drosophila Toll protein, that acts as a pathogen pattern  
CC recognition receptor, sensing the presence of conserved molecular  
CC structures present on microorganisms. The sequence was deduced  
CC from isolated cDNA clone DNA40021-1154 (ATCC 209389) (see AAX58295).  
CC The invention provides 3 novel cDNA clones that encode novel human  
CC Toll polypeptides PRO285, PRO286 (see AAY05867) and PRO358 (see  
CC AAY05868). It also provides specific antibodies and chimeric  
CC molecules comprising a PRO285, PRO286 or PRO358 polypeptide, or a  
CC transmembrane domain-deleted or inactivated variant, fused to a  
CC heterologous amino acid sequence, such as an epitope tag or  
CC immunoglobulin FC region. Being homologues of Drosophila Toll  
CC protein, the 3 human proteins are likely to be involved in adaptive  
CC immunity, particularly inflammation, septic shock, and response to  
CC pathogens in diseases aggravated by the immune response, such as  
CC diabetes, amyotrophic lateral sclerosis, cancer, rheumatoid  
CC arthritis and ulcers. The PRO polypeptides are used to identify  
CC other proteins involved in Toll-mediated signal transduction (e.g.  
CC natural ligands), to screen for receptor and ligand mimics, and to  
CC generate antibodies. Antibodies specific for the PRO polypeptides  
CC (or for the known receptor TLR-2) are used to treat septic shock  
CC (claimed).  
XX Sequence 1049 AA;

Query Match 100.0%; Score 807; DB 20; Length 1049;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPKTLPCDVTLDVPHKHHVVDCTDKHLTEIPGGIPTNTNLTLTINHPIIDISPASPHRLD 60  
DB 30 fptklpcdvtldvpkhvvdctdkhlteipggiptntnltnltinhpidispasphrid 89  
QY 61 HLVEIDFRNCVPIPIGSKNNMCIKRLQIKPRFSGLTYLKSLYLDGNOLLEIPQGLPPS 120  
DB 90 hlveidfrncvpiplgsknnmcikrlqikprfsgltylkslyldngnlleipqglpps 149  
QY 121 LQLLSLEANNIFSIKENTELANIELLYLGQNCYRNPYVSYSIEKDAFLNLTKLKVL 180  
DB 150 lqlslleannifsiikenetelanielellylgqncyrnpyvsysiekdafnltnlklvl 209  
QY 181 SKDNNVTAVPTVPLSTLTLYLNNMIKIQEDDNNLNQILDSGNCPCRYNAPFP 240  
DB 210 skdnnvtavptvplstltelyllynnmiakiqeddfnnlnqildsgncprcrynafp 269  
QY 241 CAPCKNNSLPQIPVNAFDALTEKVLRLHNSLSQHVPPRWFKNINKLQELDISQFLAKE 300  
DB 270 capcknnsplqpvnafdalteklvrlhnsnlsqhvpprwfkninklqeldisqfklake 329  
QY 301 IGDAKFLHLPSTLIQDLISFNPELOVYRASMLNSQAFSSLSKSLKILRIRCYVPKELKSN 360  
DB 330 igdakflhlpstliqdlisfnfelqvyrasmnlsqafsslskslklirirgyvfkelsfn 389  
QY 361 LSPHLNLQNLVLDLGTNFTIKANLSMFKQFKRLKVIDLSVKNISPSGDSSEVGFCSNAR 420  
DB 390 lsphlqlnlvldlgtntfikanlsmfkqfkrkvidlsvknispsgdssevgfcsnar 449  
QY 421 TSVESYEPQVLEQLHYFRYDKYARSCFRKNKEASFMSVNESCYKYGQTLDLKSNISFFVK 480

DB 450 tsvesyepqvleqlhyfrydkyarscrfknkeasfmsvnescykygqtdlslknsiffvk 509  
QY 481 SSDFOHLSFLKCLNLSGNLISOTLNGSEFOPLAEKRYLDPSNNRLDLLHSTAFEEELHKL 540  
DB 510 ssdfqnlsflkclnlsngnlisotlncsefqplaelryldfsmrldllhstafeelhk 569  
QY 541 VLDISSNSHYFQSEGITHMLNFTKNLKVLOKLMNDNDISSSTSRMTSESLRTLEFRGN 600  
DB 570 vldissnshyfseqegithmlnftknlkvlqklmndndissstsrmtseslrtllefrgn 629  
QY 601 HLDVLMREGDNRYLQFLKNLLKLELDISKNSLSFLPSGVFDGMPNPNLKNLSLAKNGLKS 660  
DB 630 hldvlmregdnrylqfknllkleeldisknslsflpsgvfdgmpnlnklsaknglks 689  
QY 661 FSWKKLQCLKNLETLDSLHNLQTLTTPERLSNCSRSLSKLNLLKNNQIRSLTKYFLQDAFOL 720  
DB 690 fswkkqlcqlknletldslshnqlttvperlsncsrslnllknnqirsltkyflqdafql 749  
QY 721 RYLDLSSNRIQIQTSTFPENVLNKLMLLLHNRFLCTCDVAFVFWVWVNHTEVTIPYLA 780  
DB 750 ryldlssnriqiqtstfpenvlnlkmlllhnrflectcdavfwfwnhntevtipyla 809  
QY 781 TDVTCVGPGAHKGQSVISLDLYTCELD 807  
DB 810 tdvtcvpggahkgqsvisdlytceid 836

RESULT 4  
AAB44323  
ID AAB44323 standard; Protein: 1049 AA.  
XX AAB44323;  
XX AAB44323;  
DT 08-FEB-2001 (first entry)  
XX Human PRO285 protein sequence SEQ ID NO:496.  
DE Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;  
KW expressed sequence tag; detection; cancer.  
XX Homo sapiens.  
OS WO2000053756-A2.  
PN 14-SEP-2000.  
XX 18-FEB-2000; 2000WO-US04341.  
XX 08-MAR-1999; 99WO-US05028.  
PR 12-MAR-1999; 99US-0123957.  
PR 29-MAR-1999; 99US-0126773.  
PR 21-APR-1999; 99US-0130232.  
PR 28-APR-1999; 99US-0131445.  
PR 14-MAY-1999; 99US-0134287.  
PR 23-JUN-1999; 99US-0141037.  
PR 26-JUL-1999; 99US-0145698.  
PR 29-OCT-1999; 99US-0162506.  
PR 30-NOV-1999; 99WO-US28313.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
XX (GETH ) GENENTECH INC.  
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;  
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;



DR N-PSDB; AAS21422.  
XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
PT other PRO polypeptides, link bioactive molecules to cells expressing  
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
PT lung, breast, prostate, cervical -  
PS Claim 12; Fig 358; 813pp; English.  
XX AAU12172-AAU12446 represent novel human secretory and transmembrane  
CC PRO polypeptides. The PRO polypeptides are useful to detect other  
CC PRO polypeptides, to link bioactive molecules to cells expressing  
CC PRO polypeptides, and to detect the presence of cells expressing  
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample.  
CC Some of the 275 sequences are also useful to stimulate the release of  
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.  
XX Sequence 1049 AA;  
SQ

Query Match  
Best Local Similarity 100.0%; Score 807; DB 22; Length 1049;  
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FPKTLPDVTLDVPHKHNIVDCTDKHLTEIPGGIPTNTNLTLTINHHPDISPASFRILD 60  
DB |||||||  
DB 30 fpktpcdvldvpknhvldctdkhlteipggiptntnltiltinhhipdispasfhrld 89  
QY 61 HLVEIDFRNCVPIPLGSKNNCKIKRLQIKPRFSGLTYLKSLYLDGQLLEIPOGLPPS 120  
DB |||||||  
DB 90 hlveidfrncvpiplgsknnckikrlqikprfsgltylkslyldgngllleipgglpps 149  
QY 121 LQLLSLEANNIFSIRKENTELANTELIYLGQNCYRNPYVSYSIEKDAFLNLTJLKVYL 180  
DB |||||||  
DB 150 lqllsleannifsirkentelanteliylgqncyrnpcyysysiekdafnltnltklvl 209  
QY 181 SLKDNNTAVPTVPLSTFELYLYNNMIAKIQEDDFNNLNQILDLSGNCPCRYNAPPP 240  
DB |||||||  
DB 210 slkdnnvtavptvplstfelylynnmiakiqeddfnnlnqildlsngcpcrynapfp 269  
QY 241 CAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLOHVPFRFKNINKLQELDSLQFLAKE 300  
DB |||||||  
DB 270 capcknnsplqipvnafdaltelkvlrlhsnslohpvrwfkinklqeldlsqflake 329  
QY 301 IGDAKFLHPLPSLIQDLDFNFELOVYRASMLLSQAFSLKSLKILIRIGYVFKELKSPN 360  
DB |||||||  
DB 330 igdakflhplpsliqdlfdfnfeolvyrasmllsqafslkslkilrirygvfkelksfn 389  
QY 361 LSPHLNLQNLVLDLGTNFIKIANLSMFQKRLKVIDLSVKNKISPSGDSSEVGFCSNAR 420  
DB |||||||  
DB 390 lsphlnlqnlvldlgtgnfikianlsmfqkrlkvidlsvknkispsgdssevgfcsnar 449  
QY 421 TSVESYEPVOLQLHYFRDYDKYARSCRFNKASPMVNESCYKQGTLDLKSNIFFVK 480  
DB |||||||  
DB 450 tsvesyepvolqlhyfrdydkyarscrfnkasmvnescykygqtdlksnsiffvk 509  
QY 481 SSDFOHLSFKLCNLISGNLISOTLNGSEFOPLAELRYLDFSNRLDLLHSTAFEELHKL 540  
DB |||||||  
DB 510 ssdfqlsflkcnlsgnlisqtlngsefqplaelryldfnsnrlldllhstafeelhkle 569

QY 541 VLDISSNSHYFQSEGITHMLNFTKNLKVLOKLMMNDNDISSSTSRMTSESLRTEFRGN 600  
DB |||||||  
DB 570 vldissnshyfqsegithmlnftknkvloklmndndissstsrmtseslrtefrgn 629  
QY 601 HLDVLRWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSCVFCGMPNPNLKNLSLAKNGLKS 660  
DB |||||||  
DB 630 hldvlrwregdnrylqlfknllkleeldisksnlsflpsgvfdgmpbnlknlsaknglks 689  
QY 661 FSWKKLQCLKNLETLDLSHNQLTTPPERLSNCSRSLSKLNLLKNNQIRSLTKYFLQDAFOL 720  
DB |||||||  
DB 690 fswkklqclknletldlshnqlttpperlsncsrslnllknmqrsrsltkyflqdafql 749  
QY 721 RYLDLSSNNKIOMIQKTSFPENVLNNLKMLLHHNRFLCTCDAVFWFVWYNNHTTEVTPYLA 780  
DB |||||||  
DB 750 ryldlssnnkiomiqktsfpenvlnlkmllhhnrflctcdavfwfwnhtevtipyla 809  
QY 781 TDVTCVGPCHAGKQSVISLDLYTCELD 807  
DB |||||||  
DB 810 tdvtcvgpahkgqgsvisldlytceld 836

RESULT 6  
AAM58342  
ID AAM58342 standard; Protein: 432 AA.  
AC AAM58342;  
XX  
DT 05-NOV-2001 (first entry)  
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 30447.  
XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX Homo sapiens.  
OS  
XX WO200157275-A2.  
PN  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-483446/52.  
XX  
DR Single exon nucleic acid probes for analyzing gene expression in human brains -  
PT  
PT  
XX Example 4; SEQ ID NO: 30447; 650pp + Sequence Listing; English.  
PS The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.  
XX  
XX Sequence 432 AA;

Query Match		50.1%;	Score 404;	DB 22;	Length 432;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 404;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	FPKTLPCDVTLOVPKNNHIVDCTDKHLTEIPGGIPTNTNLTITINHIPDISPASFHRUD 60			
Db	29	fpktpcdvtldvpknhvvdctdkhlteipggiptntnltinnhipdispasfhrld 88			
QY	61	HLVEIDFRNCVPIPLGSKNNMCKIRLQIKPRSFGLTWLKSILYLDGNQLLEIPQGLPPS 120			
Db	89	hlveidfrncvpiplgsknnmckirliqkprsfsgltylksilyldgnqlleipqglpps 148			
QY	121	LQLLSLEANNLFSIRKENITELANETILYLGQCYRNPVYSYSEKDAFLNLTKLKV 180			
Db	149	lqlisleannlfsirkenitelaneliylgqcyrnpvcyvsysiekdaflnltklkvl 208			
QY	181	SLKDNWTVAPTVLPSTLTETELYLYNNMIAKIQEDDFNNLNQIQLIDLSGNCPCYNAPFP 240			
Db	209	slkdnnvtavptvlpstltelylynnmiakiqeddfnnlnqilidsgncprcynapfp 268			
QY	241	CAPCKNNSPLQIPVNAFDALTELKVLRLHNSLIQHVPPRWEKNINKLQELDLSONFLAKE 300			
Db	269	capcknnsplqipvnafdalteklvrlhnsliqhvpprfwnkninklqeldisqnfake 328			
QY	301	IGDAKFLHFLPSLIQLDLSFNELQVYRASNNLSQAFSSLSKSLKILIRGYVFKELKSFN 360			
Db	329	igdakflhlfpслиqlildlsfnelqvyrasmllsqafsslsksklirgyvfkelksfn 388			
QY	361	LSPHLNLQNLVEVLDTGNFNIKIANLSMFQPKRLKVIDLSVNKI 404			
Db	389	lsplhnlqnlevlidgtnfikianlsmfkqfkrkvidlsvnki 432			
RESULT 7					
AAM90355					
ID	AAM90355 standard; Protein; 121 AA.				
XX					
AC	AAM90355;				
XX					
DT	07-NOV-2001 (first entry)				
XX					
DE	Human immune/haematopoietic antigen SEQ ID NO:17948.				
XX					
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;				
KW	cytostatic; gene therapy; vaccine; metastasis.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200157182-A2.				
XX					
PD	09-AUG-2001.				
XX					
PF	17-JAN-2001; 2001WO-US01354.				
XX					
PR	31-JAN-2000; 2000US-0179065.				
PR	04-FEB-2000; 2000US-0180628.				
PR	24-FEB-2000; 2000US-0184664.				
PR	02-MAR-2000; 2000US-0186350.				
PR	16-MAR-2000; 2000US-0189874.				
PR	17-MAR-2000; 2000US-0190076.				
PR	18-APR-2000; 2000US-0198123.				
PR	19-MAY-2000; 2000US-0205515.				
PR	07-JUN-2000; 2000US-0209467.				
PR	28-JUN-2000; 2000US-0214886.				
PR	30-JUN-2000; 2000US-0215135.				
PR	07-JUL-2000; 2000US-0216647.				
PR	07-JUL-2000; 2000US-0216880.				
PR	11-JUL-2000; 2000US-0217487.				
PR	11-JUL-2000; 2000US-0217496.				
PR	14-JUL-2000; 2000US-0218290.				
PR	26-JUL-2000; 2000US-0220963.				
PR	26-JUL-2000; 2000US-0220964.				
PR	14-AUG-2000; 2000US-0224518.				

PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226868.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	01-NOV-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI, INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.  
N-PSDB; AAK63136.

Nucleic acids encoding human immune/haematopoietic antigen polypeptides,  
useful for preventing, diagnosing and/or treating cancers and  
metastasis -

Claim 11: SEQ ID NO 17948; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
activity, and can be used in gene therapy and vaccine production. (I)  
proteins and polynucleotides may be used in the prevention, diagnosis and  
treatment of diseases associated with inappropriate (I) expression. For  
example, they may be used to treat disorders associated with decreased  
expression by rectifying mutations or deletions in a patient's genome  
that affect the activity of (I) by expressing inactive proteins or to  
supplement the patients own production of (I). Additionally, (I)  
polynucleotides may be used to produce the secreted (I), by inserting  
the nucleic acids into a host cell and culturing the cell to express the  
protein. (I) proteins and polynucleotides may be used to prevent,  
diagnose and treat immune/haematopoietic-related diseases, especially  
cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
to AAK87694 represent human immune/haematopoietic antigen genomic

CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
XX represent sequences used in the exemplification of the present invention.  
SQ Sequence 121 AA;  
  
Query Match 7.1%; Score 57; DB 22; Length 121;  
Best Local Similarity 100.0%; Pred. No. 5.3e-47;  
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 116 GLPPSLQLLSLEANNIFSRKENLTLANTEILYLGONCYRRNCPYVSYSIEKDAFL 172  
|||||  
Db 21 gippslqlisleannifsrkenltelanleilylgncyyrrnpcyvsysiekdafl 77  
  
RESULT 8  
AAU17407  
ID AAU17407 standard; Protein; 161 AA.  
XX  
AC AAU17407;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Novel signal transduction pathway protein, Seq ID 972.  
XX  
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
XX acquired immune deficiency syndrome.  
  
OS Homo sapiens.  
XX  
PN WO200154733-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01312.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 03-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0233168.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244517.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2000US-0254097.  
PR 05-JAN-2001; 2000US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-465460/50.  
DR N-PSDB; AAS27324.  
XX  
PT Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders -  
PS Claim 1; SEQ ID NO 972; 880pp; English.  
XX  
CC The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative  
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
CC Addison's disease), reproductive system disorders, gastrointestinal  
CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
CC as stimulators of B-cell responsiveness to pathogens, activators of  
CC T-cells, to induce higher affinity antibodies, and as a means to induce  
CC tumour proliferation in pathologies e.g. acquired immune deficiency  
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
CC pathway protein, amino acid sequences of the invention.  
XX

Query Match 5.1%; Score 41; DB 22; Length 161;



Best Local Similarity 100.0%; Pred. No. 2.4e-31; Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 767 WVNHTETIPYLATDVTCVPGAHKGQSVISLDLYTCELD 807  
Db 1 wvnhtetipylatdvtcvpgahkgqsvisdlytceld 41

RESULT 9  
AAM43641  
ID AAM43641 standard; Protein; 161 AA.  
XX AAM43641;  
AC AAM43641;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 319.  
XX  
KW Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;  
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;  
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;  
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;  
KW antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;  
KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;  
KW cardiovascular disorder; neurological disease; infection; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200155308-A2.  
XX  
PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01309.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 14-AUG-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231988.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.



PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-488781/53.  
XX N-PSDB; AAI63872.  
XX New isolated nucleic acids and polypeptides, useful for diagnosing, treating and/or preventing human diseases and disorders -  
XX Claim 11; SEQ ID NO 244; 664pp + Sequence Listing; English.  
XX The invention relates to human polynucleotides (AAI63803-AAI64012) and the encoded proteins (AAM434497-AAM43660) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes were isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.  
XX Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 235 AA;  
SQ  
Query Match 5.1%; Score 41; DB 22; Length 235;  
Best Local Similarity 100.0%; Pred. No. 3.2e-31;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 767 WVNHTVTIPYLATDVTGCGAHKGQSVISLDLYTCELD 807  
Dy 3 wvnhtvtipylatdvtcvgpghkgqsvlsldlytceld 43  
RESULT 11  
AAW86363  
ID AAW86363 standard; Protein; 394 AA.  
XX  
AC AAW86363;  
XX  
XX 15-MAR-1999 (first entry)  
XX Mouse DNAX toll-like receptor DTLR6.  
XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor; interleukin 1 receptor; phosphate metabolism; innate immunity response; modulate inflammatory function; morphological effect; immunological disorder.  
XX Mus sp.  
XX W09850547-A2.  
PN

```
XX 12-NOV-1998.
PD
XX
PF
XX 07-MAY-1998; 98WO-US08979.
XX
PR 05-MAR-1998; 98US-0076947.
PR 07-MAY-1997; 97US-0044293.
PR 22-JAN-1998; 98US-0072212.
XX
XX (SCHE ) SCHERING CORP.
XX
PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;
DR N-PSDB; AAV80677.
DR
XX
XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter
PT phosphate metabolism, modulate inflammatory function or innate
PT immunity responses
XX
PS Example; Page 154-155; 17lpp; English.
XX
CC The present invention specifically describes human DNAX toll-like
CC receptors 2 to 10 (DTLR2-10). The present sequence is mouse DTLR6
CC given in the present invention. Also described are: (1) a fusion
CC protein comprising a DTLR protein or peptide; (2) a binding compound,
CC preferably an antibody or antibody fragment which specifically binds to
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or
CC peptide; (4) an expression vector comprising the nucleic acid of (3);
CC and (5) a host cell comprising the vector of (4). The host cell of (5)
CC can be used to produce the DTLR proteins. The DTLR proteins can be used
CC to alter phosphate metabolism, to modulate inflammatory function, innate
CC immunity responses or morphological effects. The DTLR proteins can be
CC used in the treatment of conditions exhibiting abnormal expression of
CC the receptors of their ligands. These abnormalities are typically
CC manifested by immunological disorders.
XX
SQ Sequence 394 AA;

Query Match 4.2%; Score 34; DB 20; Length 394;
Best Local Similarity 100.0%; Pred. No. 3.le-24;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 774 VTIPYLATDVTGCGAHKGQSVISLDLYTCELD 807
DB 147 vtipyldtvtcvgpghkgqsvlsldlytcltd 180

RESULT 12
AAW86362
ID AAW86362 standard; Protein; 100 AA.
XX
AC AAW86362;
XX
XX 15-MAR-1999 (first entry)
DT
DE Mouse DNAX toll-like receptor DTLR6.
XX
KW DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;
KW modulate inflammatory function; morphological effect;
KW immunological disorder.
XX
OS Mus sp.
XX
XX WO9805047-A2.
PN
XX
PD 12-NOV-1998.
XX
XX 07-MAY-1998; 98WO-US08979.
PF
XX 05-MAR-1998; 98US-0076947.
PR
XX 07-MAY-1997; 97US-0044293.
PR 22-JAN-1998; 98US-0072212.
XX
XX (SCHE ) SCHERING CORP.
XX
PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;
DR N-PSDB; AAV80677.
DR
XX
XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter
PT phosphate metabolism, modulate inflammatory function or innate
PT immunity responses
XX
PS Example; Page 154-155; 17lpp; English.
XX
CC The present invention specifically describes human DNAX toll-like
CC receptors 2 to 10 (DTLR2-10). The present sequence is mouse DTLR6
CC given in the present invention. Also described are: (1) a fusion
CC protein comprising a DTLR protein or peptide; (2) a binding compound,
CC preferably an antibody or antibody fragment which specifically binds to
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or
CC peptide; (4) an expression vector comprising the nucleic acid of (3);
CC and (5) a host cell comprising the vector of (4). The host cell of (5)
CC can be used to produce the DTLR proteins. The DTLR proteins can be used
CC to alter phosphate metabolism, to modulate inflammatory function, innate
CC immunity responses or morphological effects. The DTLR proteins can be
CC used in the treatment of conditions exhibiting abnormal expression of
CC the receptors of their ligands. These abnormalities are typically
CC manifested by immunological disorders.
XX
SQ Sequence 394 AA;

Query Match 2.0%; Score 16; DB 20; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.le-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 LKVLSLKDNNTAVPT 192
DB 15 lkvislkdnntavpt 30

RESULT 13
AAG65892
ID AAG65892 standard; protein; 483 AA.
XX
AC AAG65892;
XX
XX 11-FEB-2002 (first entry)
DT
DE Amino acid sequence of GSK gene Id 90060.
XX
XX Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;
XX antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;
XX cytostatic; cerebroprotective; vasotropic; human.
XX
OS Homo sapiens.
XX
XX WO200172961-A2.
PN
XX
XX 04-OCT-2001.
PD
XX
XX 22-MAR-2001; 2001WO-US09226.
PF
XX
XX 24-MAR-2000; 2000US-192158P.
PR 28-MAR-2000; 2000US-192668P.
PR 27-APR-2000; 2000US-200166P.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
PA
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;
```

PI Lai Y;  
 XX WPI: 2001-639223/73.  
 DR N-PSDB; AAI67182.  
 XX  
 XX Isolated polypeptides, which may be peptide hormones, which are  
 PT identified by high throughput genome-based biology which identifies  
 PT genes and gene products as therapeutic targets for treatment of  
 PT diseases such as diabetes and cancer -  
 XX  
 XX Claim 1; Page 65-66; 99pp; English.  
 PS  
 XX  
 CC The invention provides polypeptides (AAG65886-65918) which may be peptide  
 CC hormones (including insulin, growth hormones, chemokines, cytokines,  
 CC neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic  
 CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
 CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
 CC high throughput genome-based biology and polynucleotides (AAI67176-67208)  
 CC encoding them. The polypeptides can be expressed by standard recombinant  
 CC methodology. The polypeptides are useful in the treatment of disease such  
 CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
 CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
 CC asthma, manic depression, dementia, delirium, mental retardation,  
 CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
 CC or sexual development disorders, and dysfunctions of the blood cascade  
 CC system including those leading to stroke. The polynucleotides may be used  
 CC as diagnostic reagents through detecting mutations in the associated gene  
 CC and for chromosome localization and for tissue expression studies. The  
 CC polypeptides and polynucleotides may also be used as vaccines.  
 XX  
 XX Sequence 483 AA;  
 SQ

Query Match 1.5%; Score 12; DB 22; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 0.0088;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQLTTPV 686  
 |||||  
 Db 173 ldlshnqlttvp 184

RESULT 14  
 AAG65893  
 ID AAG65893 standard; protein; 605 AA.  
 XX  
 AC AAG65893;  
 XX  
 XX 11-FEB-2002 (first entry)  
 DT  
 XX  
 DE Amino acid sequence of GSK gene Id 90060.  
 XX  
 KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
 KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
 KW cytotstatic; cerebroprotective; vasotropic; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200172961-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 XX 22-MAR-2001; 2001WO-US09226.  
 PF  
 XX  
 XX 24-MAR-2000; 2000US-192158P.  
 PR  
 PR 28-MAR-2000; 2000US-192668P.  
 PR  
 PR 27-APR-2000; 2000US-200166P.  
 XX  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 PI  
 PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
 PI Lai Y;

XX WPI: 2001-639223/73.  
 DR N-PSDB; AAI67183.  
 XX  
 XX Isolated polypeptides, which may be peptide hormones, which are  
 PT identified by high throughput genome-based biology which identifies  
 PT genes and gene products as therapeutic targets for treatment of  
 PT diseases such as diabetes and cancer -  
 XX  
 XX Claim 1; Page 67-68; 99pp; English.  
 PS  
 XX  
 CC The invention provides polypeptides (AAG65886-65918) which may be peptide  
 CC hormones (including insulin, growth hormones, chemokines, cytokines,  
 CC neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic  
 CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
 CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
 CC high throughput genome-based biology and polynucleotides (AAI67176-67208)  
 CC encoding them. The polypeptides can be expressed by standard recombinant  
 CC methodology. The polypeptides are useful in the treatment of disease such  
 CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
 CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
 CC asthma, manic depression, dementia, delirium, mental retardation,  
 CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
 CC or sexual development disorders, and dysfunctions of the blood cascade  
 CC system including those leading to stroke. The polynucleotides may be used  
 CC as diagnostic reagents through detecting mutations in the associated gene  
 CC and for chromosome localization and for tissue expression studies. The  
 CC polypeptides and polynucleotides may also be used as vaccines.  
 XX  
 XX Sequence 605 AA;  
 SQ

Query Match 1.5%; Score 12; DB 22; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQLTTPV 686  
 |||||  
 Db 295 ldlshnqlttvp 306

RESULT 15  
 AAE07271  
 ID AAE07271 standard; peptide; 14 AA.  
 XX  
 AC AAE07271;  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX  
 DE Human leucine-rich repeat (LRR) signature sequence #6.  
 XX  
 KW Human; leucine-rich repeat protein-like; LRR; cytostatic; gene therapy;  
 KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
 KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
 KW angiolasty-related restenosis; viral infection; rheumatoid arthritis;  
 KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
 KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
 KW parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
 KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
 KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157261-A1.  
 XX  
 PD 09-AUG-2001.  
 XX  
 XX 02-FEB-2001; 2001WO-US03653.  
 PF  
 XX  
 XX 03-FEB-2000; 2000US-0496914.  
 PR  
 PR 27-APR-2000; 2000US-0560875.  
 PR  
 PR 27-SEP-2000; 2000US-0672221.  
 XX



CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
CC The present sequence is human gap protein used in the exemplification  
CC of the invention.  
XX  
SQ Sequence 121 AA:  
  
Query Match 1.4%; Score 11; DB 22; Length 121;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 672 LETLDLSHNQL 682  
Db 91 letldlshnql 101  
|||||  
letldlshnql 101  
  
RESULT 17  
AAE07278  
ID AAE07278 standard; Protein: 674 AA.  
XX  
AC AAE07278;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human leucine-rich repeat (LRR) protein #1 fragment.  
XX  
KW Human; leucine-rich repeat protein-like; LRR; cytostatic; gene therapy;  
KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW angiotensin-related stenosis; viral infection; rheumatoid arthritis;  
KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157261-A1.  
XX  
PD 09-AUG-2001.  
XX  
PF 02-FEB-2001; 2001WO-US03653.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
XX  
PR 27-APR-2000; 2000US-0560875.  
XX  
PR 27-SEP-2000; 2000US-0672221.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
PI Drmanac RT, Wang M, Chen L, Yang Y;  
XX  
DR WPI; 2001-496930/54.  
XX  
Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
PT for diagnosing, treating bleeding disorders, myocardial infarction,  
PT atherosclerosis, angiotensin-related stenosis and glomerular diseases  
PT -  
XX  
Claim 12; Page 137-139; 156pp; English.  
XX  
The present sequence is a fragment of human leucine-rich repeat  
CC (LRR) protein. LRR protein is involved in protein recognition, cell  
CC adhesion, development, signal transduction, DNA repair, recombination,  
CC immune responses and transcription. LRR DNA and protein are useful for  
CC treating, preventing haemophilia, bleeding disorders (Bernard-Soulier  
CC syndrome), myocardial infarction, thrombosis, atherosclerosis, glomerular  
CC diseases, angiotensin-related stenosis, viral infections, melanomas,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,

CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
CC irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX  
SQ Sequence 674 AA:  
  
Query Match 1.4%; Score 11; DB 22; Length 674;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 672 LETLDLSHNQL 682  
Db 234 letldlshnql 244  
|||||  
letldlshnql 244  
  
RESULT 18  
AAE07281  
ID AAE07281 standard; Protein: 674 AA.  
XX  
AC AAE07281;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human leucine-rich repeat (LRR) protein #2 fragment.  
XX  
KW Human; leucine-rich repeat protein-like; LRR; cytostatic; gene therapy;  
KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW angiotensin-related stenosis; viral infection; rheumatoid arthritis;  
KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157261-A1.  
XX  
PD 09-AUG-2001.  
XX  
PF 02-FEB-2001; 2001WO-US03653.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
XX  
PR 27-APR-2000; 2000US-0560875.  
XX  
PR 27-SEP-2000; 2000US-0672221.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
PI Drmanac RT, Wang M, Chen L, Yang Y;  
XX  
DR WPI; 2001-496930/54.  
XX  
Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
PT for diagnosing, treating bleeding disorders, myocardial infarction,  
PT atherosclerosis, angiotensin-related stenosis and glomerular diseases  
PT -  
XX  
Claim 12; Page 151-153; 156pp; English.

XX The present sequence is a fragment of human leucine-rich repeat  
CC (LRR) protein. LRR protein is involved in protein recognition, cell  
CC adhesion, development, signal transduction, DNA repair, recombination,  
CC immune responses and transcription. LRR DNA and protein are useful for  
CC treating, preventing haemophilia, bleeding disorders (Bernard-Soulier  
CC syndrome), myocardial infarction, thrombosis, atherosclerosis, glomerular  
CC diseases, angiotensin-related stenosis, viral infections, melanomas,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
CC irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock.  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX  
SQ Sequence 674 AA;

Query Match 1.4%; Score 11; DB 22; Length 674;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 LETLDSHNL 682  
|||||||  
Db 234 letldishnql 244

RESULT 19  
AAU29303  
ID AAU29303 standard; Protein; 692 AA.

XX AC AAU29303;  
XX  
XX 18-DEC-2001 (first entry)  
XX  
XX Human PRO polypeptide sequence #280.

XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
XX  
OS Homo sapiens.

XX WO200168848-A2.

XX 20-SEP-2001.

XX 28-FEB-2001; 2001WO-US06520.

XX 01-MAR-2000; 2000WO-US05601.

PR 02-MAR-2000; 2000WO-US05841.

PR 03-MAR-2000; 2000US-187202P.

PR 06-MAR-2000; 2000US-186968P.

PR 14-MAR-2000; 2000US-189320P.

PR 14-MAR-2000; 2000US-189328P.

PR 15-MAR-2000; 2000WO-US06884.

PR 21-MAR-2000; 2000US-190828P.

PR 21-MAR-2000; 2000US-191007P.

PR 21-MAR-2000; 2000US-191048P.

PR 21-MAR-2000; 2000US-191314P.

PR 28-MAR-2000; 2000US-192655P.

PR 29-MAR-2000; 2000US-193032P.  
PR 29-MAR-2000; 2000US-193053P.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 04-APR-2000; 2000US-194449P.  
PR 04-APR-2000; 2000US-194647P.  
PR 11-APR-2000; 2000US-195975P.  
PR 11-APR-2000; 2000US-196000P.  
PR 11-APR-2000; 2000US-196187P.  
PR 11-APR-2000; 2000US-196690P.  
PR 11-APR-2000; 2000US-196820P.  
PR 18-APR-2000; 2000US-198121P.  
PR 18-APR-2000; 2000US-198585P.  
PR 25-APR-2000; 2000US-199397P.  
PR 25-APR-2000; 2000US-199550P.  
PR 03-MAY-2000; 2000US-201516P.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 05-JUN-2000; 2000US-209832P.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 22-AUG-2000; 2000US-0644848.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
XX  
XX (GETH ) GENENTECH INC.

PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

XX WPI; 2001-602746/68.

DR N-PSDB; AAS46204.

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the

PT presence of tumours, such as prostate and breast tumours, in mammals and  
PT to screen for modulators of the compounds -  
XX Claim 11; Fig 560; 774pp; English.

PS Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

CC The PRO polypeptides and their associated nucleic acids can be used to  
CC detect the presence of a tumour in a mammal by comparing the level of  
CC expression of a PRO polypeptide in a test sample of cells from the animal  
CC and a control sample of normal cells, whereby a higher level of  
CC expression in the test sample indicates the presence of a tumour in the  
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats  
CC and rabbits but are preferably human. The polypeptides can be used to  
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,  
CC when contacted with it. A specific polypeptide can be used to stimulate  
CC the proliferation or differentiation of chondrocyte cells. The PRO  
CC proteins can be used to determine the presence of tumours and also  
CC susceptibility to tumour development, particularly adrenal, lung, colon,  
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian  
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
CC can be used for genetic analysis of individuals with genetic disorders.

XX SQ Sequence 692 AA;

Query Match 1.4%; Score 11; DB 22; Length 692;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 LETLDSHNL 682

|||||||

Db 252 letldishnql 262

RESULT 20

AAE07266



ID AAE07266 standard; Protein; 692 AA.  
AC AAE07266;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human leucine-rich repeat (LRR) protein #1.  
XX  
KW Human; leucine-rich repeat protein-like; LRR; cytosolic; gene therapy;  
KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW angioplasty-related restenosis; viral infection; rheumatoid arthritis;  
KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein /label= Signal\_peptide  
FT Region 19..692 /label= Mature\_LRR\_protein  
FT Region 134..147  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 180..193  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 225..238  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 252..265  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 351..364  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 375..388  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 378..391  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 535..548  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Region 560..573  
FT Region /label= Leucine\_rich\_repeat\_signature  
FT Domain 648..673  
FT /label= Transmembrane\_domain  
XX  
XX WO200157261-A1.  
XX  
XX 09-AUG-2001.  
XX  
XX 02-FEB-2001; 2001WO-US03653.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
XX PR 27-APR-2000; 2000US-0560875.  
XX PR 27-SEP-2000; 2000US-0672221.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
XX Drmanac RT, Wang M, Chen L, Yang Y,  
XX  
XX WPI; 2001-496930/54.  
XX DR N-PSDB; AAD13550.  
XX  
XX Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
XX for diagnosing, treating bleeding disorders, myocardial infarction,  
XX atherosclerosis, angioplasty-related restenosis and glomerular diseases  
XX  
XX Claim 12; Page 131-133; 156pp; English.  
XX  
XX The present sequence is human leucine-rich repeat (LRR) protein.  
XX LRR protein is involved in protein recognition, cell adhesion,  
CC

CC development, signal transduction, DNA repair, recombination, immune  
CC responses and transcription. LRR DNA and protein are useful for treating,  
CC preventing haemophilia, bleeding disorders (Bernard-Soulier syndrome),  
CC myocardial infarction, thrombosis, atherosclerosis, glomerular diseases,  
CC angioplasty-related restenosis, viral infections, melanomas,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
CC irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX  
SQ Sequence 692 AA;  
  
Query Match 1.4%; Score 11; DB 22; Length 692;  
Best Local Similarity 100.0%; Pred. No. 0.11;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 672 LETDLSHNL 682  
Db 252 letldlshnql 262  
|||||  
letldlshnql 262  
  
RESULT 21  
AAE07280  
ID AAE07280 standard; Protein; 692 AA.  
XX  
AC AAE07280;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human leucine-rich repeat (LRR) protein #2.  
XX  
KW Human; leucine-rich repeat protein-like; LRR; cytosolic; gene therapy;  
KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW angioplasty-related restenosis; viral infection; rheumatoid arthritis;  
KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein /label= Signal\_peptide  
FT Region 19..692 /label= Mature\_LRR\_protein  
FT Region 134..147 /label= Leucine\_rich\_repeat\_signature  
FT Region 180..193 /label= Leucine\_rich\_repeat\_signature  
FT Region 225..238 /label= Leucine\_rich\_repeat\_signature  
FT Region 252..265 /label= Leucine\_rich\_repeat\_signature  
FT Region 351..364 /label= Leucine\_rich\_repeat\_signature  
FT Region 375..388 /label= Leucine\_rich\_repeat\_signature

FT                    /label= Leucine\_rich\_repeat\_signature  
 FT                    378..391  
 FT                    /label= Leucine\_rich\_repeat\_signature  
 FT                    535..548  
 FT                    /label= Leucine\_rich\_repeat\_signature  
 FT                    560..573  
 FT                    /label= Leucine\_rich\_repeat\_signature  
 FT                    648..673  
 FT                    /label= Transmembrane\_domain  
 XX  
 PN                    WO200157261-A1.  
 XX  
 PD                    09-AUG-2001.  
 XX  
 PF                    02-FEB-2001; 2001WO-US03653.  
 XX  
 PR                    03-FEB-2000; 2000US-0496914.  
 PR                    27-APR-2000; 2000US-0568875.  
 XX                    27-SEP-2000; 2000US-0672221.  
 XX  
 PA                    (HYSE-) HYSEQ INC.  
 XX  
 PI                    Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
 PI                    Drmanac RT, Wang M, Chen L, Yang Y;  
 XX  
 DR                    WPI; 2001-496930/54.  
 DR                    N-PSDB; AAD13552.  
 XX  
 PT                    Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
 PT                    for diagnosing, treating bleeding disorders, myocardial infarction,  
 PT                    atherosclerosis, angioplasty-related restenosis and glomerular diseases  
 PT                    .  
 XX  
 PS                    Claim 12; Page 149-151; 156pp; English.  
 XX  
 CC                    The present sequence is human leucine-rich repeat (LRR) protein.  
 CC                    LRR protein is involved in protein recognition, cell adhesion,  
 CC                    development, signal transduction, DNA repair, recombination, immune  
 CC                    responses and transcription. LRR DNA and protein are useful for treating,  
 CC                    preventing haemophilia, bleeding disorders (Bernard-Soulier syndrome),  
 CC                    myocardial infarction, thrombosis, atherosclerosis, glomerular diseases,  
 CC                    angioplasty-related restenosis, viral infections, melanomas,  
 CC                    immunological disorders (rheumatoid arthritis, multiple sclerosis,  
 CC                    psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
 CC                    periodontitis); wound healing, burns, ulcers, incisions and cancer.  
 CC                    LRR is also useful for proliferation of neural cells and nerve  
 CC                    regeneration, for treating peripheral nervous system diseases, central  
 CC                    nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
 CC                    disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
 CC                    mechanical and traumatic disorders (spinal cord disorders, head trauma)  
 CC                    cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
 CC                    irritation associated with infection (septic shock, sepsis or systemic  
 CC                    inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
 CC                    endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
 CC                    nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
 CC                    anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
 CC                    chronic inflammatory arthritis, pancreatic cell damage from diabetes  
 CC                    mellitus type 1, graft versus host disease, inflammation associated with  
 CC                    pulmonary disease, other autoimmune diseases or inflammatory diseases.  
 XX  
 SQ                    Sequence    692 AA;

Query Match                    1.4%;    Score 11;    DB 22;    Length 692;  
 Best Local Similarity    100.0%;    Pred. No. 0.11;    Mismatches    0;    Gaps    0;  
 Matches    11;    Conservative    0;    Indels    0;    Gaps    0;

QY    672 LETLDLSHNL 682  
 DB    252 letldlshnl 262  
 IIIIIIIIII

RESULT    22

AAW86356  
 ID    AAW86356 standard; Protein; 329 AA.  
 XX  
 AC    AAW86356;  
 XX  
 DT    15-MAR-1999    (first entry)  
 XX  
 DE    Partial human DNAX toll-like receptor DTLR7.  
 XX  
 KW    DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;  
 KW    interleukin 1 receptor; phosphate metabolism; innate immunity response;  
 KW    modulate inflammatory function; morphological effect;  
 KW    immunological disorder.  
 XX  
 OS    Homo sapiens.  
 XX  
 PN    WO9850547-A2.  
 XX  
 PD    12-NOV-1998.  
 XX  
 PF    07-MAY-1998;    98WO-US08979.  
 XX  
 PR    05-MAR-1998;    98US-0076947.  
 PR    07-MAY-1997;    97US-0044293.  
 PR    22-JAN-1998;    98US-0072212.  
 XX  
 PA    (SCHE ) SCHERING CORP.  
 XX  
 PI    Bazan JF, Hardiman GT, Kastelein RA, Rock FL;  
 XX  
 DR    WPI; 1999-059670/05.  
 DR    N-PSDB; AAV80670.  
 XX  
 PT    Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter  
 PT    phosphate metabolism, modulate inflammatory function or innate  
 PT    immunity responses  
 XX  
 PS    Claim 6; Page 131-132; 171pp; English.  
 XX  
 CC    The present invention specifically describes human DNAX toll-like  
 CC    receptors 2 to 10 (DTLR2-10). The present sequence is partial human  
 CC    DTLR7 given in the present invention. Also described are: (1) a fusion  
 CC    protein comprising a DTLR protein or peptide; (2) a binding compound,  
 CC    preferably an antibody or antibody fragment which specifically binds to  
 CC    a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
 CC    peptide; (4) an expression vector comprising the nucleic acid of (3);  
 CC    and (5) a host cell comprising the vector of (4). The host cell of (5)  
 CC    can be used to produce the DTLR proteins. The DTLR proteins can be used  
 CC    to alter phosphate metabolism, to modulate inflammatory function, innate  
 CC    immunity responses or morphological effects. The DTLR proteins can be  
 CC    used in the treatment of conditions exhibiting abnormal expression of  
 CC    the receptors of their ligands. These abnormalities are typically  
 CC    manifested by immunological disorders.  
 XX  
 SQ                    Sequence    329 AA;

Query Match                    1.2%;    Score 10;    DB 20;    Length 329;  
 Best Local Similarity    100.0%;    Pred. No. 0.56;    Mismatches    0;    Gaps    0;  
 Matches    10;    Conservative    0;    Indels    0;    Gaps    0;

QY    225 LDLSGNCPRC 234  
 DB    86 ldlsgncprc 95  
 IIIIIIIIIII

RESULT    23  
 AAB64892  
 ID    AAB64892 standard; Protein; 426 AA.  
 XX  
 AC    AAB64892;  
 XX  
 DT    23-MAR-2001    (first entry)

XX Human secreted protein sequence encoded by gene 11 SEQ ID NO:70.  
DE  
XX  
DE Human: secreted protein; diagnosis; immunomodulatory; antisclerotic;  
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;  
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;  
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vulneryary;  
KW antialzheimers; antiparkinsonian; antimicrobial; immune disorder;  
KW multiple sclerosis; systemic lupus erythematosus; HIV; infection;  
KW hyperproliferative disorder; cancer; Gaucher's disease; wound healing;  
KW cardiovascular disease; Scimitar syndrome; Chaga's cardiomyopathy;  
KW coronary arteriosclerosis; angiogenic disorder; diabetic retinopathy;  
KW corneal graft neovascularisation; neurological disorder; regeneration;  
KW Huntington's chorea; Alzheimer's disease; Parkinson's disease;  
KW infectious disease; chemotaxis.  
XX  
OS Homo sapiens.  
XX  
XX WO200076530-A1.  
PN  
XX  
PD 21-DEC-2000.  
XX  
XX 01-JUN-2000; 2000WO-US14933.  
PF  
XX  
XX 11-JUN-1999; 99US-0138572.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
PA  
XX  
XX Rosen CA, Ruben SM, Komatsoulis GA;  
PI  
XX  
XX WPI; 2001-071147/08.  
DR  
XX N-PSDB; AAF33223.  
DR  
XX  
XX Nucleic acids encoding 49 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
PT disease and diabetic retinopathy -  
PT  
XX  
PS Claim 11; Page 489-490; 554pp; English.  
XX  
XX The polynucleotide sequences given in AAF33213 to AAF33261 encode the  
CC human secreted proteins given in AAB64882 to AAB64930. AAB64931 to  
CC AAB64991 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Examples of activities include:  
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;  
CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;  
CC vascular; antimicrobial; anti-angiogenic; ophthalmological;  
CC neuroprotectant; anticonvulsant; nootropic; antialzheimers;  
CC antiparkinsonian; and vulneryary. The polynucleotides and polypeptides can  
CC be used in the prevention, diagnosis and treatment of diseases associated  
CC with inappropriate polypeptide expression. Disorders that may be  
CC prevented, diagnosed and/or treated by the above methods include immune  
CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and  
CC human immuno-deficiency virus (HIV) infections), hyperproliferative  
CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases  
CC (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary  
CC arteriosclerosis), angiogenic disorders (e.g. corneal graft  
CC neovascularisation and diabetic retinopathy), neurological disorders  
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
CC infectious diseases and/or for promoting wound healing, regeneration and  
CC /or chemotaxis. AAF33204 to AAF33212 and AAB64881 represent sequences  
CC used in the exemplification of the present invention.  
XX  
XX Sequence 426 AA;

Query Match 1.2%; Score 10; DB 22; Length 426;  
Best Local Similarity 100.0%; Pred. No. 0.69;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 LETLDLSHNO 681

Db 252 letldlshng 261  
|||||||  
RESULT 24  
AAU14800  
ID AAU14800 standard; Protein; 504 AA.  
XX  
AC AAU14800;  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Novel bone marrow polypeptide #6.  
XX  
KW Bone marrow; immunosuppressive; immunostimulant; cytostatic; vulneryary;  
KW nootropic; neuroprotective; therapeutic; antigenic; nutritional source;  
KW cytokine; stem cell growth factor; tissue regeneration; cancer;  
KW Parkinson's disease; Alzheimer's disease; neurodegenerative disorder;  
KW wound healing; immune system; autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200155442-A2.  
XX  
PD 02-AUG-2001.  
XX  
XX 25-JAN-2001; 2001WO-US02543.  
PF  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
PR 30-NOV-2000; 2000US-0250583.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Zhang J;  
PI Drmanac RT;  
PI  
XX  
XX WPI; 2001-465578/50.  
DR N-PSDB; AAS23105.  
DR  
XX  
XX Novel polynucleotides encoding bone marrow-derived polypeptides useful  
PT for treating, e.g., cancer, autoimmune disease and Alzheimer's disease  
PT  
XX  
XX Claim 10; Page 173-174; 274pp; English.  
PS  
XX  
XX AAU14795-AAU14973 represent the amino acid sequences of novel bone  
CC marrow-derived polypeptides. The proteins may exhibit e.g., cytokine or  
CC stem cell growth factor activity and may be useful for re-engineering  
CC damaged or diseased tissues, producing large quantities of human cells to  
CC treat Parkinson's, Alzheimer's and other neurodegenerative diseases,  
CC wound healing, immune system stimulation or suppression, treating  
CC autoimmune diseases, and cancer. The corresponding nucleic acid sequences  
CC can be used to express recombinant protein for analysis, characterisation  
CC or therapeutic use; as markers for tissues in which the corresponding  
CC protein is preferentially expressed; as a molecular weight marker on  
CC gels; as chromosome markers or tags; as probes to hybridise and discover  
CC novel, related DNA sequences; as a source of information to derive  
CC polymerase chain reaction (PCR) primers; for selecting and making  
CC oligomers for attachment to a 'gene chip' or other support; to raise  
CC anti-protein antibodies using DNA immunisation techniques; and as an  
CC antigen to raise anti-DNA antibodies or elicit another immune response.  
CC The proteins may be also used as a nutritional source.  
XX  
XX Sequence 504 AA;

Query Match 1.2%; Score 10; DB 22; Length 504;  
Best Local Similarity 100.0%; Pred. No. 0.8;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 LDLSGNCPC 234  
|||||  
Db 154 ldlsgncpc 163

## RESULT 25

ABB23027  
ID ABB23027 standard; Protein; 1040 AA.

AC ABB23027;

XX 23-JAN-2002 (first entry)

DE Protein #5026 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human hearts -

XX Claim 15; SEQ ID No 24797; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABA21533-ABA41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1040 AA;

Query Match 1.2%; Score 10; DB 22; Length 1040;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 LDLSGNCPC 234

|||||

Db 250 ldlsgncpc 259

## RESULT 26

AAM31141

ID AAM31141 standard; Protein; 1040 AA.

XX AAM31141;

XX 17-OCT-2001 (first entry)

XX Peptide #5178 encoded by probe for measuring placental gene expression.

DE Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 31410; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs; see AA131315-AA157546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.

XX Sequence 1040 AA;

Query Match 1.2%; Score 10; DB 22; Length 1040;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 LDLSGNCPC 234

|||||

Db 250 ldlsgncpc 259

## RESULT 27

AAY41768

ID AAY41768 standard; Protein; 1041 AA.

XX AAY41768;

XX 07-DEC-1999 (first entry)

XX Human PRO286 protein sequence.

XX Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;  
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;  
KW secreted protein; transmembrane protein.

XX Homo sapiens.

XX

PN WO9946281-A2.  
XX 16-SEP-1999.  
XX 08-MAR-1999; 99WO-US05028.  
XX 10-MAR-1998; 98US-0077450.  
PR 11-MAR-1998; 98US-0077632.  
PR 11-MAR-1998; 98US-0077641.  
PR 11-MAR-1998; 98US-0077649.  
PR 12-MAR-1998; 98US-0077791.  
PR 13-MAR-1998; 98US-0078004.  
PR 17-MAR-1998; 98US-0040220.  
PR 20-MAR-1998; 98US-0078886.  
PR 20-MAR-1998; 98US-0078910.  
PR 20-MAR-1998; 98US-0078936.  
PR 20-MAR-1998; 98US-0078939.  
PR 25-MAR-1998; 98US-0079294.  
PR 26-MAR-1998; 98US-0079556.  
PR 27-MAR-1998; 98US-0079663.  
PR 27-MAR-1998; 98US-0079664.  
PR 27-MAR-1998; 98US-0079689.  
PR 27-MAR-1998; 98US-0079728.  
PR 27-MAR-1998; 98US-0079786.  
PR 30-MAR-1998; 98US-0079920.  
PR 30-MAR-1998; 98US-0079923.  
PR 31-MAR-1998; 98US-0080105.  
PR 31-MAR-1998; 98US-0080107.  
PR 31-MAR-1998; 98US-0080165.  
PR 31-MAR-1998; 98US-0080194.  
PR 01-APR-1998; 98US-0080327.  
PR 01-APR-1998; 98US-0080328.  
PR 01-APR-1998; 98US-0080333.  
PR 01-APR-1998; 98US-0080334.  
PR 08-APR-1998; 98US-0081049.  
PR 08-APR-1998; 98US-0081070.  
PR 08-APR-1998; 98US-0081071.  
PR 09-APR-1998; 98US-0081195.  
PR 09-APR-1998; 98US-0081203.  
PR 09-APR-1998; 98US-0081229.  
PR 15-APR-1998; 98US-0081817.  
PR 15-APR-1998; 98US-0081838.  
PR 15-APR-1998; 98US-0081952.  
PR 15-APR-1998; 98US-0081955.  
PR 21-APR-1998; 98US-0082568.  
PR 21-APR-1998; 98US-0082569.  
PR 22-APR-1998; 98US-0082700.  
PR 22-APR-1998; 98US-0082704.  
PR 22-APR-1998; 98US-0082804.  
PR 23-APR-1998; 98US-0082767.  
PR 23-APR-1998; 98US-0082796.  
PR 27-APR-1998; 98US-0083336.  
PR 28-APR-1998; 98US-0083322.  
PR 29-APR-1998; 98US-0083392.  
PR 29-APR-1998; 98US-0083495.  
PR 29-APR-1998; 98US-0083496.  
PR 29-APR-1998; 98US-0083499.  
PR 29-APR-1998; 98US-0083500.  
PR 29-APR-1998; 98US-0083545.  
PR 29-APR-1998; 98US-0083554.  
PR 29-APR-1998; 98US-0083558.  
PR 29-APR-1998; 98US-0083559.  
PR 30-APR-1998; 98US-0083742.  
PR 05-MAY-1998; 98US-0084366.  
PR 06-MAY-1998; 98US-0084414.  
PR 06-MAY-1998; 98US-0084441.  
PR 07-MAY-1998; 98US-0084598.  
PR 07-MAY-1998; 98US-0084600.  
PR 07-MAY-1998; 98US-0084627.  
PR 07-MAY-1998; 98US-0084637.  
PR 07-MAY-1998; 98US-0084639.  
PR 07-MAY-1998; 98US-0084640.  
PR 07-MAY-1998; 98US-0084643.

PR 13-MAY-1998; 98US-0085323.  
PR 13-MAY-1998; 98US-0085338.  
PR 13-MAY-1998; 98US-0085339.  
PR 15-MAY-1998; 98US-0085573.  
PR 15-MAY-1998; 98US-0085579.  
PR 15-MAY-1998; 98US-0085580.  
PR 15-MAY-1998; 98US-0085582.  
PR 15-MAY-1998; 98US-0085689.  
PR 15-MAY-1998; 98US-0085697.  
PR 15-MAY-1998; 98US-0085700.  
PR 15-MAY-1998; 98US-0085704.  
PR 18-MAY-1998; 98US-0086023.  
PR 22-MAY-1998; 98US-0086392.  
PR 22-MAY-1998; 98US-0086414.  
PR 22-MAY-1998; 98US-0086430.  
PR 22-MAY-1998; 98US-0086486.  
PR 28-MAY-1998; 98US-0087098.  
PR 28-MAY-1998; 98US-0087106.  
PR 28-MAY-1998; 98US-0087208.  
PR 30-JUL-1998; 98US-0094651.  
PR 11-SEP-1998; 98US-0100038.  
XX (GETH ) GENENTECH INC.  
PA Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;  
XX WPI; 1999-551358/46.  
XX N-PSDB; AA234304.  
PT New secreted and transmembrane polypeptides and their polynucleotides,  
PT useful for treating blood coagulation disorders, cancers and cellular  
PT adhesion disorders.  
XX Claim 12; Fig 211; 530pp; English.  
XX The present invention describes secreted and transmembrane polypeptides  
CC and their polynucleotides. The nucleotide sequences are useful as  
CC sources of probes, primers, for chromosome mapping, and for generation  
CC of antisense sequences. They can also be used to create transgenic  
CC animals. The proteins can be used to treat a variety of diseases and  
CC disorders, depending on their function. Diseases that may be treated  
CC include blood coagulation disorders, cancers and cellular adhesion  
CC disorders. They may also be used to raise antibodies. AA233891 to  
CC AA234338, and AA41685 to AA41774 represent polynucleotide and  
CC polypeptide sequence given in the exemplification of the present  
CC invention.  
XX Sequence 1041 AA;  
SQ

Query Match 1.2%; Score 10; DB 20; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 225 LDLSGNCPRC 234  
| | | | |  
Db 251 ldlsncprc 260

RESULT 28  
AA05867  
ID AA05867 standard; Protein; 1041 AA.  
XX  
AC AA05867;  
XX  
DT 02-AUG-1999 (first entry)  
XX  
DE Human Toll protein PRO286.  
XX  
KW PRO286; Toll; homologue; human; adaptive immunity; septic shock;  
KW inflammation; diabetes; amyotrophic lateral sclerosis; cancer;  
KW ulcer; rheumatoid arthritis; pathogen pattern recognition receptor;  
KW signal transduction.



PR 08-MAR-1999; 99WO-US05028.  
 PR 12-MAR-1999; 99US-0123957.  
 PR 29-MAR-1999; 99US-0126773.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 29-OCT-1999; 99US-0162506.  
 PR 30-NOV-1999; 99WO-US28313.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 30-DEC-1999; 99WO-US31274.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 XX PA (GETH ) GENENTECH INC.  
 XX PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gershtsen ME;  
 PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
 PI Kijavirij IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;  
 PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;  
 DR WPI: 2000-611443/58.  
 DR N-PSDB; AAC78584.  
 XX Novel PRO polypeptides and polynucleotides used in detection methods,  
 PT to target bioactive molecules to specific cells, and to modulate  
 PT cellular activities -  
 XX Claim 12; Fig 211; 636pp; English.  
 XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed  
 CC sequence tag) sequences which encode secreted or transmembrane PRO  
 CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic  
 CC activity. The polynucleotides and polypeptides can be used for detecting  
 CC the presence of PRO polypeptides in samples, for linking bioactive  
 CC molecules to cells and for modulating biological activities of cells,  
 CC using the polypeptides for specific targeting. The polypeptide targeting  
 CC can be used to kill the target cells, e.g. for the treatment of cancers.  
 CC The polypeptide pairs provide specific targeting of bioactive molecules  
 CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in  
 CC the isolation of the PRO polynucleotide sequences.  
 XX SQ Sequence 1041 AA;  
 Query Match 1.2%; Score 10; DB 21; Length 1041;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 225 LDIISGNCPRC 234  
 Db |||||  
 251 ldlsgncprc 260  
 RESULT 30  
 AAR13130  
 ID AAR13130 standard; Protein; 15 AA.  
 XX AC AAR13130;  
 XX DT 01-OCT-1991 (first entry)  
 XX GPIB alpha peptide fragment.  
 KW Von Willebrand factor; vWF; platelet membrane glycoprotein Ib;  
 KW glycoalbumin; thrombosis.  
 XX

OS Synthetic.  
 XX WO9109614-A.  
 PN 11-JUL-1991.  
 PD 04-JAN-1991; 91WO-US00087.  
 PF 14-NOV-1990; 90US-0613083.  
 PR 04-JAN-1990; 90US-0460674.  
 XX (SCRI-) SCRIPPS CLINIC & RE.  
 PA Ruggeri ZM, Zimmerman TS, Houghten RA, Vicente V, Mohri H;  
 PI Ware JL;  
 XX WPI: 1991-222654/30.  
 DR GPIB alpha peptide fragment - inhibits binding of von Willebrand  
 PT factor to platelet membrane glyco-protein Ib, useful in treating  
 PT thrombosis.  
 XX Claim 1; Page 56; 76pp; English.  
 PS The peptide corresponds to residues 81-95 of the N-terminus of  
 CC glycoalbumin, a water sol. proteolytic fragment of GPIB alpha. It  
 CC may be linked to a second peptide from the 45 kD N-terminal  
 CC tryptic fragment of GPIB alpha. The peptide inhibits binding of  
 CC vWF to GPIB. It can be used to inhibit activation, aggregation  
 CC and/or adhesion of platelets, esp. for inhibition of thrombosis.  
 CC See also AAR13128-R13138.  
 XX SQ Sequence 15 AA;  
 Query Match 1.1%; Score 9; DB 12; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.39;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 674 TLDLSHNL 682  
 Db |||||  
 1 tldlshnl 9  
 RESULT 31  
 AAM40831  
 ID AAM40831 standard; Protein; 114 AA.  
 XX AC AAM40831;  
 XX DT 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 5762.  
 DE Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX OS Homo sapiens.  
 XX PN WO200153312-A1.  
 XX PD 26-JUL-2001.  
 XX 26-DEC-2000; 2000WO-US34263.  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR N-PSDB; AAI59987.  
 XX  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 2; SEQ ID NO 5762; 10078pp; English.  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic, and  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 SQ Sequence 114 AA;  
 Query Match 1.1%; Score 9; DB 22; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 722 YLDLSSNKI 730  
 Db 105 yldlssnki 113  
 RESULT 32  
 AAB44116  
 ID AAB44116 standard; Protein; 155 AA.  
 XX  
 AC AAB44116;  
 XX  
 XX 08-FEB-2001 (first entry)  
 DT  
 XX Human cancer associated protein sequence SEQ ID NO:1561.  
 DE  
 XX Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;  
 KW antidiabetic; antithrombotic; antirheumatic; antiarthritic; antiviral;  
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; neurotropic;  
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200055350-A1.  
 PN  
 XX

PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05882.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 PI WPI; 2000-587533/55.  
 XX N-PSDB; AAC78325.  
 DR  
 DR Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer -  
 XX  
 PS Claim 11; Page 2238-2239; 2352pp; English.  
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB4398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnerary; immunomodulator;  
 CC antidiabetic; antithrombotic; antirheumatic; antiarthritic;  
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC neurotropic; vasotropic; antipsoriatic and antiangiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention.  
 XX  
 SQ Sequence 155 AA;  
 Query Match 1.1%; Score 9; DB 21; Length 155;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 113 IPQGLPPSL 121  
 Db 60 ipqglppsl 68  
 RESULT 33  
 AAY94900  
 ID AAY94900 standard; Protein; 158 AA.  
 XX  
 AC AAY94900;  
 XX  
 XX 16-JUN-2000 (first entry)  
 DT  
 XX Human secreted protein clone ns197\_1 protein sequence SEQ ID NO:6.  
 DE  
 XX Human; secreted protein; immunestimulant; immunosuppressant; virucide;  
 KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;  
 KW antidiabetic; antithrombotic; antirheumatic; antidiabetic; antirheumatic;  
 KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;  
 KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;  
 KW connective tissue disease; multiple sclerosis; erythematosis;  
 KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;  
 KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
 KW insulin dependent diabetes mellitus; graft-versus-host-disease;  
 KW autoimmune inflammatory eye disease; allergy.  
 XX





Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 LETLDSHN 680  
|||||  
Db 219 letldishn 227

RESULT 35  
AAR42265  
ID AAR42265 standard; Protein; 234 AA.  
XX  
AC AAR42265;  
XX  
DT 28-APR-1994 (first entry)  
XX  
DE Decorin sequence PT-76 (N-terminal to LRR8).

KW leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;  
KW fusion protein; maltose binding protein; tumour growth; inhibition;  
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.

XX WO9320202-A.  
XX 14-OCT-1993.  
XX 02-APR-1993; 93WO-US03171.  
XX 03-APR-1992; 92US-0865652.  
XX (LJOL-) LA JOLLA CANCER RES FOUND.

PA Cardenas J, Craig W, Mullen DG, Pierschbacher MD;

PI Ruoslahti EI;

XX WPI; 1993-336010/42.

DR N-PSDB; AAQ50051.

XX Active fragments of protein esp. decorin - with cell regulatory  
PT factor domain, useful for inhibiting cell regulatory factor  
PT activity

XX Claim 10; Page 45-46; 77pp; English.

XX Active fragments of decorin (full-length coding sequence AAQ500046)  
CC were generated by PCR and fused to Maltose Binding Protein. The  
CC resulting fusion proteins were useful for inhibiting the activity of  
CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
CC conditions associated with over-activity of the growth factor such  
CC as certain tumours.

XX Sequence 234 AA;

Query Match 1.1%; Score 9; DB 14; Length 234;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121  
|||||  
Db 187 ipqglppsl 195

RESULT 36  
AAU20516  
ID AAU20516 standard; Protein; 235 AA.  
XX  
AC AAU20516;  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Human secreted protein, Seq ID No 508.  
XX  
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;

KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytostatic; Alzheimer's disease; Parkinson's disease; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
XX thrombosis; wound healing.

OS Homo sapiens.

PN WO200155326-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01347.

PR 31-JAN-2000; 2000US-0179065.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

DR WPI; 2001-451931/48.

DR N-PSDB; AAS33225.

XX New nucleic acids and polypeptides, useful for diagnosing, preventing  
or treating medical conditions -

PS Claim 11; SEQ ID No 508; 753pp; English.

XX The invention relates to novel isolated nucleic acid molecules (I)  
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
CC the prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression of secreted proteins. (I) and complementary  
CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
CC of similar nucleic acid sequences in samples, and so which patients may  
CC be in need of restorative therapy. (II) may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of the secreted  
CC proteins. The anti-(II) antibodies and antagonists may also be used to  
CC down regulate expression and activity of (II). The anti-(II) antibodies  
CC may also be used as diagnostic agents for detecting the presence of (II)  
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/

CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,  
CC angina and thrombosis), infections caused by bacteria, viruses and  
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),  
CC agonists, antagonists and antibodies can also be used to promote wound  
CC healing, maintain organs before transplantation, and support cell culture  
CC of primary tissues. AAU20342-AAU20665 represent human secreted protein  
CC amino acid sequences, and related sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO  
CC at: ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 235 AA;

Query Match 1.1%; Score 9; DB 22; Length 235;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 YLDLSSNKI 730  
|||||  
Db 82 yidlssnki 90

```
RESULT 37
AAR42266
ID AAR42266 standard; Protein; 280 AA.
XX
AC AAR42266;
XX
DT 28-APR-1994 (first entry)
XX
DE Decorin sequence PT-77 (N-terminal to LRR10).
XX
DE leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;
KW fusion protein; maltose binding protein; tumour growth; inhibition;
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.
XX
PN WO9320202-A.
XX
PD 14-OCT-1993.
XX
PF 02-APR-1993; 93WO-US03171.
XX
PR 03-APR-1992; 92US-0865652.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI Cardenas J, Craig W, Mullen DG, Pierschbacher MD;
PI Ruoslahti EI;
XX
DR WPI; 1993-336910/42.
DR N-PSDB; AAQ50052.
XX
PT Active fragments of protein esp. decorin - with cell regulatory
PT factor domain, useful for inhibiting cell regulatory factor
activity
XX
PS Claim 10; Page 47-48; 77pp; English.
XX
CC Active fragments of decorin (full-length coding sequence AAQ50046)
CC were generated by PCR and fused to Maltose Binding Protein. The
CC resulting fusion proteins were useful for inhibiting the activity of
CC a cell regulatory factor, esp. TGF-beta, and hence for treating
CC conditions associated with over-activity of the growth factor such
CC as certain tumours.
XX
SQ Sequence 280 AA;

Query Match 1.1%; Score 9; DB 14; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121
DB 187 IPQGLPPSL 195

RESULT 38
AAP91368
ID AAP91368 standard; peptide; 293 AA.
XX
AC AAP91368;
XX
DT 21-MAR-1990 (first entry)
XX
DE 45 kDa amino terminal tryptic fragment of glycoocalicin and derivs.
XX
KW Glycoocalicin; von Willebrand factor; platelet membrane glycoprotein 1b;
KW platelet aggregation prevention; thrombosis inhibition; antithrombotic
KW agent
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
```

```
FT Peptide 18..34
FT Peptide 21..35
FT Peptide 26..40
FT Peptide 26..34
FT Peptide 141..155
FT Peptide 231..245
FT Peptide 271..285
FT Peptide 281..285
XX
PN EP317278-A.
XX
PD 24-MAY-1989.
XX
PF 16-NOV-1988; 88EP-0310799.
XX
PR 17-NOV-1987; 87US-0121454.
XX
PA (SCRI-) SCRIPPS CLINIC RES.
XX
PI Zimmerman TS, Ruggeri ZM, Houghten RA, Vincete V, Mohri H;
XX
DR WPI; 1989-152756/21.
XX
PT Proteolytic 45 KD fragment of glycoocalicin and derivs. - which inhibit
PT binding of von Willebrand factor to platelet membrane glyco:protein, and
PT used as antithrombotic agents
XX
PS Claim 1; ; 10pp; English.
XX
CC The claim is for a peptide of a 45kD amino terminal tryptic fragment of
CC glycoocalicin selected from the sequence which inhibits binding of von
CC Willebrand factor to platelet membrane glycoprotein 1b and related
CC molecules or other cells and cell matrices. Also claimed are a sequential
CC subset of the above (Claim 2) and specific peptides (see FT) (Claim 3)
CC with the same functions and any peptide of any sequential subset of amino
CC acids of the sequence (Claim 4). The peptides and derivs. prevent
CC platelet aggregation and inhibit thrombosis.
XX
SQ Sequence 293 AA;

Query Match 1.1%; Score 9; DB 10; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNQL 682
DB 81 TLDLSHNQL 89

RESULT 39
AAR42267
ID AAR42267 standard; Protein; 305 AA.
XX
AC AAR42267;
XX
DT 28-APR-1994 (first entry)
XX
DE Decorin sequence PT-78 (N-terminal to half C-terminal).
XX
KW leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;
KW fusion protein; maltose binding protein; tumour growth; inhibition;
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.
XX
PN WO9320202-A.
XX
PD 14-OCT-1993.
XX
PF 02-APR-1993; 93WO-US03171.
XX
PR 03-APR-1992; 92US-0865652.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
```

XX Cardenas J, Craig W, Mullen DG, Pierschbacher MD;  
PI Ruoslahti EI;  
XX  
XX WPI: 1993-336910/42.  
DR N-PSDB; AAQ50053.  
XX  
XX Active fragments of protein esp. decorin - with cell regulatory  
PT factor domain, useful for inhibiting cell regulatory factor  
PT activity  
XX  
XX Claim 10; Page 49-50; 77pp; English.  
XX  
XX Active fragments of decorin (full-length coding sequence AAQ50046)  
CC were generated by PCR and fused to Maltose Binding Protein. The  
CC resulting fusion proteins were useful for inhibiting the activity of  
CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
CC conditions associated with over-activity of the growth factor such  
CC as certain tumours.  
XX  
XX Sequence 305 AA;  
SQ  
  
Query Match 1.1%; Score 9; DB 14; Length 305;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 113 IPQGLPPSL 121  
Db 187 ipqglppsl 195  
|||||  
  
RESULT 40  
AAR42260  
ID AAR42260 standard; Protein; 331 AA.  
XX  
AC AAR42260;  
XX  
DT 28-APR-1994 (first entry)  
XX  
DE Mature decorin PR-65.  
XX  
XX leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;  
KW fusion protein; maltose binding protein; tumour growth; inhibition;  
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.  
XX  
FH Key Location/Qualifiers  
FT Region 1..45  
FT /label= N-terminal\_region  
FT /note= "contains 4 Cys residues"  
FT Region 46..280  
FT /label= repeat\_region  
FT /note= "contains 10 leucine-rich repeats"  
FT Region 281..331  
FT /label= C-terminal\_region  
XX  
XX WO9320202-A.  
PN  
XX  
XX 14-OCT-1993.  
PD  
XX  
XX 02-APR-1993; 93WO-US03171.  
PF  
XX  
XX 03-APR-1992; 92US-0865652.  
PR  
XX  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
PA  
XX  
XX Cardenas J, Craig W, Mullen DG, Pierschbacher MD;  
PI Ruoslahti EI;  
XX  
XX WPI: 1993-336910/42.  
DR N-PSDB; AAQ50046.  
XX  
XX Active fragments of protein esp. decorin - with cell regulatory

PT factor domain, useful for inhibiting cell regulatory factor  
PT activity  
XX  
XX Claim 10; Page 36-38; 77pp; English.  
XX  
XX Active fragments of decorin (full-length coding sequence AAQ50046)  
CC were generated by PCR and fused to Maltose Binding Protein. The  
CC resulting fusion proteins were useful for inhibiting the activity of  
CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
CC conditions associated with over-activity of the growth factor such  
CC as certain tumours.  
XX  
XX Sequence 331 AA;  
SQ  
  
Query Match 1.1%; Score 9; DB 14; Length 331;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 113 IPQGLPPSL 121  
Db 187 ipqglppsl 195  
|||||  
  
RESULT 41  
AAR89439  
ID AAR89439 standard; Protein; 342 AA.  
XX  
AC AAR89439;  
XX  
DT 20-AUG-1996 (first entry)  
XX  
DE Human recombinant decorin.  
XX  
KW Decorin; PG-II; PG-40; proteoglycan; guanidinium ion.  
XX  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..14  
FT /label= Sig\_peptide  
XX  
XX WO9601842-A1.  
PN  
XX  
XX 25-JAN-1996.  
PD  
XX  
XX 07-JUL-1995; 95WO-US08542.  
PF  
XX  
XX 08-JUL-1994; 94US-0272919.  
PR  
XX  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
PA  
XX  
XX Craig WS, Harper JR, Hernandez SD, Kostel PJ, Parker JR;  
PI Vedvick TS;  
PI  
XX  
XX WPI: 1996-097586/10.  
DR N-PSDB; AAT10741.  
DR  
XX  
XX Purificn. of human recombinant decorin - using a strong anion  
PT exchange resin, a hydrophobic interaction chromatography resin and a  
PT strong anion exchange resin  
XX  
XX Disclosure; Fig 1A-D; 55pp; English.  
PS  
XX  
XX Human recombinant decorin (AAR89439) was obtd. by expression of a  
CC cDNA clone (AAT10741) in CHO host cells. Decorin (or PGII or PG-40)  
CC is a proteoglycan having a 40 kDa core protein. Recombinant  
CC decorin can be produced by cotransfection of CHO-DC44 cells with  
CC pSV2-decorin and pSV2dhfr. Large-scale cultures can be performed  
CC using CHO cells attached to microcarrier beads. The recombinant  
CC protein is purified from the cells using a 3-step chromatographic  
CC procedure. It can be used for the highly sensitive detection of  
CC guanidinium ions (ppm range), partic. in protein-contg. solns.

CC purified using GuHCl, and also has therapeutic applns.

Query Match 1.18; Score 9; DB 17; Length 342;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ Sequence 342 AA;

Query Match 1.18; Score 9; DB 17; Length 342;

Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ Sequence 342 AA;

Qy 113 IPOGLPPSL 121

Db 199 ipqglppsl 207

RESULT 42

AAY57079

ID AAY57079 standard; protein; 359 AA.

XX

AC AAY57079;

DT 28-FEB-2000 (first entry)

DE Human decorin amino acid sequence.

XX

XX Neglected target tissue antigen; NNTA; autoimmunity; autoimmune response;

KW immunotherapeutic agent; insulin dependent diabetes mellitus;

KW multiple sclerosis; autoimmune thyroiditis; rheumatoid arthritis;

KW uveoretinitis; inflammatory response.

XX

OS Homo sapiens.

XX

XX WO9956763-A1.

PN 11-NOV-1999.

XX

XX 07-MAY-1999; 99WO-US10250.

XX

PR 07-MAY-1998; 98US-0084636.

XX

XX (REGC ) UNIV CALIFORNIA.

XX

XX Kaufman DL, Tian J, Olcott A;

PI WPI; 2000-052905/04.

XX

XX Administration of neglected target tissue antigens to modulate immune

PT responses

XX

XX Disclosure; Page 26; 79pp; English.

XX

XX Amino acid sequences AAY57063-Y57091 are examples of neglected target

CC tissue antigens NNTAs. NNTAs are antigens (whole antigens or fragments)

CC not involved in autoimmunity. These peptides and proteins are used in

CC the method of the invention which involves administering an NNTA as an

CC antigen based immunotherapeutic agent, to a host afflicted with an

CC autoimmune response associated with an autoimmune disease. The

CC immunotherapeutic agent is used to treat autoimmune diseases such as

CC insulin dependent diabetes mellitus, multiple sclerosis, autoimmune

CC thyroiditis, uveoretinitis, rheumatoid arthritis or abnormal

CC inflammatory immune responses. The NNTA induces regulatory tolerance by

CC elicitation of regulatory T cells among T cells recognizing the NNTA but

CC not participating in the immune response. The NNTA are capable of

CC recognition by substantial populations of uncommitted T cells which can

CC be primed, or biased, towards regulatory responses to provide effective

CC treatment. The NNTA are effective in regulating undesirable immune

CC responses even when target determinants used as agents promoting

CC tolerance agents have failed to induce an effective regulatory T cell

CC response. NNTAs as agents promoting tolerance are anticipated to be safer

XX than use of target determinants.

XX

SQ Sequence 359 AA;

Query Match 1.18; Score 9; DB 21; Length 359;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IPOGLPPSL 121

Db 215 ipqglppsl 223

RESULT 43

AAG78511

ID AAG78511 standard; protein; 359 AA.

XX

AC AAG78511;

DT 29-JAN-2002 (first entry)

DE Human decorin amino acid sequence.

XX

XX Fibromodulin; Decorin; Biglycan; Proteoglycan; Extracellular matrix;

KW TGF-beta; Transforming growth factor-beta; Burn; Wound; Cytostatic;

KW Nephrotropic; Antirheumatic; Antiarthritic; Vasotropic; Vulnery;

KW Antiartherosclerotic; Hepatotropic; Cardiant; Dermatological;

KW Glomerulonephritis; Rheumatoid arthritis; Arteriosclerosis;

KW Adult respiratory distress syndrome; Cirrhosis; Cancer; Fibrotic.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Peptide 1..30

FT /label= Signal\_peptide

FT Protein 31..359

FT /label= Mature\_human\_decorin

XX

XX US6277812-B1.

XX

XX 21-AUG-2001.

XX

XX 02-JUN-1995; 95US-0458834.

XX

XX 14-NOV-1991; 91US-0792192.

XX

XX 17-NOV-1992; 92US-0978931.

XX

XX 08-SEP-1994; 94US-0303238.

XX

XX 28-JUN-1988; 88US-0212702.

XX

XX 22-JAN-1990; 90US-0467888.

XX

XX 13-MAY-1992; 92US-0882345.

XX

XX (BURN-) BURNHAM INST.

XX

XX Ruoslahti EI, Yamaguchi Y;

PI WPI; 2001-610491/70.

XX

XX Use of decorin or biglycan polypeptide for inhibiting TGF-beta activity

CC in the treatment of dermal wounds and cancer

XX

XX Example 8; Fig 11; 40pp; English.

XX

XX The invention relates to the inhibition of transforming growth

CC factor-beta (TGF-beta) activity involving contacting TGF-beta with a

CC purified polypeptide comprising leucine-rich amino acid sequence of a

CC member of decorin superfamily of mammalian proteoglycans. The following

CC activities can be attributed to the polypeptide of the invention:

CC cyostatic, nephrotropic, antirheumatic, antiarthritic, vasotropic,

CC antiarteriosclerotic, hepatotropic, cardiant, dermatological and

CC vulnery. Polypeptides of the invention act as transforming growth

CC factor-beta (TGF-beta) binder. The polypeptides of the invention can be

CC used for treating a pathology, particularly proliferative pathology

CC caused by a transforming growth factor-beta (TGF-beta) regulated

CC activity such as cancer; particularly fibrotic cancer, fibrotic

CC disease, glomerulonephritis, rheumatoid arthritis, arteriosclerosis,

CC adult respiratory distress syndrome, cirrhosis of liver, fibrosis of

CC lungs, post-myocardial infarction, cardiac fibrosis, post-angioplasty

CC restenosis, renal interstitial fibrosis and certain dermal fibrotic  
 CC conditions such as keloids and scarring resulting from burn injuries;  
 CC other invasive skin injuries and reconstructive surgery. The wounds  
 CC treated with the polypeptide, particularly decorin exhibit no detectable  
 CC scarring, and are histologically normal. The current sequence represents  
 CC human decorin.

XX Sequence 359 AA;

Query Match 1.1%; Score 9; DB 22; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 5.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121

Db 215 ipqglppsl 223

RESULT 44  
 AAB49650  
 ID AAB49650 standard; Protein; 493 AA.

XX AAB49650;

XX 20-MAR-2001 (first entry)

XX Human SEC2 protein sequence SEQ ID 4.

DE  
 KW SECX; secreted protein; cancer; angiogenesis; wound healing;  
 KW immune disorder; neurodegenerative disease; allergic reaction;  
 KW respiratory problem; organ transplantation; contraceptive; human;  
 KW chromosome 1; proliferative disorder.

XX Homo sapiens.

XX WO200070046-A2.

XX 23-NOV-2000.

XX 12-MAY-2000; 2000WO-US13291.

XX 14-MAY-1999; 98US-0134315.

PR 12-JAN-2000; 2000US-0175744.

PR 10-MAR-2000; 2000US-0188274.

PA (CURA-) CURAGEN CORP.

XX Shinkets RA, Fernandes E, Boldog F;

DR WPI: 2001-025020/03.

DR N-PSDB; AAF23411.

XX New SECX polypeptides and nucleic acids useful for treating or

PT preventing cancer, other disorders related to angiogenesis,

PT neurodegenerative diseases, autoimmune disorders and allergic reactions

PT -

XX Claim 1; Page 12-14; 132pp; English.

XX Polynucleotide sequences AAF23410 - AAF23419 encode secreted SECX  
 CC proteins AAB49649 - AAB49658. Sequences AAF23420 - AAF23450 represent  
 CC primers and probes used in the isolation and characterization of the SECX  
 CC DNA sequences of the invention. The new polypeptides and nucleic acids  
 CC can be used in screening assays, detection assays, preventive or  
 CC predictive medicine, therapeutic and prophylactic treatment, and  
 CC pharmacogenomics. Specifically, the SECX polypeptides and nucleic acids  
 CC are useful for treating cancer; other disorders related to angiogenesis  
 CC e.g. abnormal wound healing, psoriasis; neurodegenerative diseases;  
 CC immune disorders; liver cirrhosis; benign tumours; fibrocystic conditions  
 CC and tissue hypertrophy (e.g. benign prostatic hypertrophy); allergic  
 CC reactions and conditions such as asthma and other respiratory problems;  
 CC as well as in treating or preventing diseases associated with organ

CC transplantation, atherosclerosis-associated diseases or disorders. The  
 CC polypeptides can also be used for bone, cartilage, tendon, ligament  
 CC and/or tissue growth or regeneration, wound healing, tissue repair and  
 CC replacement, gut protection or regeneration, as a contraceptive, to  
 CC inhibit thromboses, infections caused by bacteria, virus, fungi and other  
 CC parasites, and as a vaccine. SECX antibodies may be used to isolate or  
 CC detect SECX proteins, monitor protein level in tissue as part of a  
 CC clinical testing procedure, treat proliferative disorders including  
 CC tumours and benign hyperplasias.

XX Sequence 493 AA;

Query Match 1.1%; Score 9; DB 22; Length 493;  
 Best Local Similarity 100.0%; Pred. No. 7.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 720 LRYLDLSSN 728

Db 112 lryldlssn 120

RESULT 45

AAM39045

ID AAM39045 standard; Protein; 557 AA.

XX AAM39045;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2190.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

DR N-PSDB; AAI58201.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 CC such as central nervous system injuries -

XX Example 4; SEQ ID NO 2190; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ

Sequence 557 AA;

Query Match 1.1%; Score 9; DB 22; Length 557;  
Best Local Similarity 100.0%; Pred. No. 8.1; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 722 YLDLSSNKI 730  
|||||||  
DB 78 yldlssnki 86

## RESULT 46

AAAY49933  
ID AAY49933 standard; Protein; 562 AA.

XX

XX

AC AAY49933;

XX

DT 01-FEB-2000 (first entry)

XX

DE Human glycoprotein Ib/mouse IgG1Fc chimeric protein.

XX

XX

KW Glycoprotein Ib; glycoallidin; detection; antithrombotic; binding;

KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;

KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.

XX

OS Chimeric - Homo sapiens.

OS Chimeric - Mus musculus.

XX

FH Key Location/Qualifiers

FT Peptide 1..16

FT /label= signal

FT Protein 17..562

FT /label= human\_glycoprotein\_Ib/mouse\_IgG1Fc\_chimeric

FT protein

XX

PN WO9954360-A1.

XX

XX

PD 28-OCT-1999.

XX

XX

PF 13-JAN-1999; 99WO-JP000089.

XX

XX

PR 23-APR-1998; 98JP-0113962.

XX

XX

PA (AJIN ) AJINOMOTO CO INC.

XX

XX

PI Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;

PI Tanaka A, Shinozaki J, Jojima Y;

XX

XX

DR WPI; 2000-013233/01.

DR N-PSDB; AA235701.

XX

XX

CC A method has been developed for the detection of binding between the von  
CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in  
CC which the von Willebrand factor immobilised in a reactor reacts with  
CC glycoprotein Ib in the presence of a binding inducer to promote binding  
CC between the von Willebrand factor and glycoprotein Ib. This method is  
CC for the detection of glycoallidin as a means of thrombotic disease  
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also  
CC for screening substances with anti-thrombotic activity for the  
CC prevention and treatment of thrombotic diseases. The method is direct,  
CC convenient and quantitative, with reproducibility, and there is no need  
CC to construct a monoclonal antibody for the assay. The present sequence  
CC is a human glycoprotein Ib/mouse immunoglobulin gamma 1 Fc chimeric  
CC protein from the present invention.  
XX  
SQ

Sequence 562 AA;

Query Match 1.1%; Score 9; DB 21; Length 562;  
Best Local Similarity 100.0%; Pred. No. 8.2; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 674 TLDLSHNQL 682  
|||||||  
DB 97 tldlshnql 105

## RESULT 47

AAAY49935  
ID AAY49935 standard; Protein; 568 AA.

XX

XX

AC AAY49935;

XX

DT 01-FEB-2000 (first entry)

XX

DE Human glycoprotein Ib/mouse IgG2aFc chimeric protein #2.

XX

XX

KW Glycoprotein Ib; glycoallidin; detection; antithrombotic; binding;

KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;

KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.

XX

OS Chimeric - Homo sapiens.

OS Chimeric - Mus musculus.

XX

FH Key Location/Qualifiers

FT Peptide 1..16

FT /label= signal

FT Protein 17..568

FT /label= Human\_glycoprotein\_Ib/mouse\_IgG2aFc\_chimeric

FT protein

XX

PN WO9954360-A1.

XX

XX

PD 28-OCT-1999.

XX

XX

PF 13-JAN-1999; 99WO-JP000089.

XX

XX

PR 23-APR-1998; 98JP-0113962.

XX

XX

PA (AJIN ) AJINOMOTO CO INC.

XX

XX

PI Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;

PI Tanaka A, Shinozaki J, Jojima Y;

XX

XX

DR WPI; 2000-013233/01.

DR N-PSDB; AA235706.

XX

XX

CC Direct quantitative detection of glycoallidin with immobilized von  
CC Willebrand factor to bond with chimeric protein via inhibiting  
CC glycoprotein Ib binding, for diagnosis of thrombotic diseases and  
CC screening anti-thrombotic substances  
XX  
PS Example 1; Page 77-79; 83pp; Japanese.  
XX

CC A method has been developed for the detection of binding between the von  
 CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in  
 CC which the von Willebrand factor immobilised in a reactor reacts with  
 CC glycoprotein Ib in the presence of a binding inducer to promote binding  
 CC between the von Willebrand factor and glycoprotein Ib. This method is  
 CC for the detection of glycoalkaloid as a means of thrombotic disease  
 CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also  
 CC for screening substances with anti-thrombotic activity for the  
 CC prevention and treatment of thrombotic diseases. The method is direct,  
 CC convenient and quantitative, with reproducibility, and there is no need  
 CC to construct a monoclonal antibody for the assay. The present sequence  
 CC represents a human glycoprotein Ib/mouse immunoglobulin gamma 2a Fc  
 CC chimeric protein from the present invention.

XX Sequence 568 AA;

Query Match 1.1%; Score 9; DB 21; Length 568;  
 Best Local Similarity 100.0%; Pred. No. 8.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNL 682  
 DB 97 tldlsnql 105  
 |||||

RESULT 48  
 AAR51116  
 ID AAR51116 standard; protein; 610 AA.

XX AC AAR51116;

XX DT 23-SEP-1994 (first entry)

XX DE Platelet glycoprotein Ib alpha.

XX KW Platelet; glycoprotein; imaging; thrombolytic agent;  
 KW tissue plasminogen activator; tPA; pro-urokinase; urokinase;  
 KW streptokinase; Bernard-Soulier disease; thrombus; aggregation;  
 KW anisoylated plasminogen-streptokinase activator complex; adhesion;  
 KW inhibition.

XX OS Homo sapiens.

XX PN US5298239-A.

XX PD 29-MAR-1994.

XX PF 07-OCT-1991; 91US-0770968.

XX PR 07-OCT-1991; 91US-0770968.

XX PR 15-JAN-1992; 92US-0821717.

XX PA (UUNY ) UNIV NEW YORK STATE RES FOUND.

XX PI Cunningham D, Finch CN, Lyle VA, Miller JL;

XX DR WPI; 1994-100287/12.

XX PT Platelet glycoprotein Ib alpha with an amino acid substatn at  
 PT position 57 - has reduced reactivity with Von Willebrand factor,  
 PT and can be used to inhibit platelet aggregation and inhibition

XX PS Claim 1; Columns 3-8; 20pp; English.

XX CC A substitution in platelet glycoprotein Ib alpha (Leucine 57 to  
 CC Phenylalanine) underlies a form of Bernard-Soulier disease. The  
 CC mutated glycoprotein can be used in compositions to inhibit  
 CC platelet aggregation/adhesion. The glycoprotein may be labelled and  
 CC used as an imaging agent and may also be bound to a thrombolytic  
 CC agent, preferably tissue plasminogen activator (tPA),  
 CC (pro) urokinase, streptokinase, anisoylated plasminogen-streptokinase  
 CC activator complex, tPA analogues or a protease, allowing localisation

CC of the thrombolytic agent to a thrombus.  
 XX SQ Sequence 610 AA;

Query Match 1.1%; Score 9; DB 15; Length 610;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNL 682  
 DB 81 tldlsnql 89  
 |||||

RESULT 49  
 AAR56664  
 ID AAR56664 standard; protein; 610 AA.

XX AC AAR56664;

XX DT 23-SEP-1994 (first entry)

XX DE Mutant platelet glycoprotein Ib alpha.

XX KW Platelet; glycoprotein; imaging; thrombolytic agent;  
 KW tissue plasminogen activator; tPA; pro-urokinase; urokinase;  
 KW streptokinase; Bernard-Soulier disease; thrombus; aggregation;  
 KW anisoylated plasminogen-streptokinase activator complex; adhesion;  
 KW inhibition.

XX OS Homo sapiens.

XX PN US5298239-A.

XX PD 29-MAR-1994.

XX PF 07-OCT-1991; 91US-0770968.

XX PR 07-OCT-1991; 91US-0770968.

XX PR 15-JAN-1992; 92US-0821717.

XX PA (UUNY ) UNIV NEW YORK STATE RES FOUND.

XX PI Cunningham D, Finch CN, Lyle VA, Miller JL;

XX DR WPI; 1994-100287/12.

XX PT Platelet glycoprotein Ib alpha with an amino acid substatn at  
 PT position 57 - has reduced reactivity with Von Willebrand factor,  
 PT and can be used to inhibit platelet aggregation and inhibition

XX PS Claim 1; 20pp; English.

XX CC A substitution in platelet glycoprotein Ib alpha (Leucine 57 to  
 CC Phenylalanine) underlies a form of Bernard-Soulier disease. The  
 CC mutated glycoprotein can be used in compositions to inhibit  
 CC platelet aggregation/adhesion. The glycoprotein may be labelled and  
 CC used as an imaging agent and may also be bound to a thrombolytic  
 CC agent, preferably tissue plasminogen activator (tPA),  
 CC (pro) urokinase, streptokinase, anisoylated plasminogen-streptokinase  
 CC activator complex, tPA analogues or a protease, allowing localisation  
 CC of the thrombolytic agent to a thrombus.

XX SQ Sequence 610 AA;

Query Match 1.1%; Score 9; DB 15; Length 610;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNL 682  
 DB 81 tldlsnql 89  
 |||||



Search completed: July 17, 2002, 09:46:12  
Job time: 124 sec

```
RESULT 50
AAR89436
ID AAR89436 standard; Protein; 610 AA.
XX
AC AAR89436;
XX
DT 02-SEP-1996 (first entry)
XX
DE Mutated platelet glycoprotein-Ib-alpha GPIbA protein sequence.
XX
KW Platelet glycoprotein-Ib-alpha; GPIbA; mutagenesis; point mutation;
von Willebrand factor; blood disorder; platelet disorder;
XX protein engineering; Bernard-Soulier disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 57 /note= "substitution from Phe in wild-typ
FT GPIb-alpha"
FT Region 36..200 /note= "Leu rich area"
FT Region 200..220 /note= "Leu rich area"
FT Region 220..310 /note= "flanking region to Leu rich area"
FT Region 310..420 /note= "hinge region"
FT Region 310..420 /note= "Ser/Thr rich area"
XX
US5492809-A.
PN
XX
XX 20-FEB-1996.
XX
XX 07-OCT-1991; 91US-0770968.
XX
PR 15-JAN-1992; 92US-0821717.
PR 07-OCT-1991; 91US-0770968.
PR 09-SEP-1993; 93US-0119262.
XX
PA (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX
PI Cunningham D, Finch CN, Lyle VA, Miller JL;
XX
WPI; 1996-128585/13.
XX
PT DNA encoding platelet glyco:protein Ib alpha mutant Phe57 -
PT introduced into platelets to reduce aggregation and reactivity with
PT von Willebrand factor, also probe for diagnosis of Bernard-Soulier
PT disease
XX
PS Disclosure; Column 21-26; 20pp; English.
XX
CC A substitution of T for C at position 259 in the DNA sequence of
CC GPIb-alpha leads to the replacement of Phe for Leu at
CC residue 57 of the mature GPIb-alpha molecule. This mutated GPIb-
CC alpha protein is less reactive with von Willebrand factor, a
CC a characteristic of the autosomal recessive bleeding disorder,
CC Bernard-Soulier disease.
XX
SQ Sequence 610 AA;

Query Match 1.1%; Score 9; DB 17; Length 610;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNQL 682
Db 81 tldlsnql 89
```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2002, 09:44:08 ; Search time 23.61 Seconds  
(without alignments)  
3284.375 Million cell updates/sec

Title: US-09-202-054-1\_COPY\_30\_836

Perfect score: 807

Sequence: 1 FPKTLPCDVTLDVKNHVIV.....PQAHKQSVISLDLYTCELD 807

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 6

Total number of hits satisfying chosen parameters: 3826

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	1.5	1232	2 T05322	hypothetical prote
2	10	1.2	367	2 AC1328	internalin protein
3	10	1.2	791	2 S67265	hypothetical prote
4	9	1.1	359	1 NBHUC8	decorin precursor
5	9	1.1	360	2 S06280	decorin - rabbit
6	9	1.1	360	2 I47020	platelet glycoprot
7	9	1.1	626	1 NBHUIA	chaoptin precursor
8	9	1.1	1134	1 A29944	flightless-I homol
9	9	1.1	1268	2 A49674	hypothetical prote
10	9	1.1	1770	2 S56221	hypothetical prote
11	9	1.1	1784	2 C96615	hypothetical prote
12	8	1.0	283	2 SI0005	pyridoxal kinase (
13	8	1.0	288	2 AI0810	hypothetical prote
14	8	1.0	371	2 T49908	hypothetical prote
15	8	1.0	506	2 T51702	amidophosphoribos
16	8	1.0	537	2 S26857	isocitrate lyase (
17	8	1.0	607	2 E96598	protein F20N2.2 [1
18	8	1.0	612	2 T10727	protein kinase Xa2
19	8	1.0	636	2 S47299	gene F protein - r
20	8	1.0	743	2 C84633	probable disease r
21	8	1.0	768	2 T17462	disease resistance
22	8	1.0	783	2 T45899	receptor protein k
23	8	1.0	806	2 A53256	nuclear protein bi
24	8	1.0	838	2 A96557	probable receptor
25	8	1.0	853	2 T17461	disease resistance
26	8	1.0	855	2 T07015	Cf-4A protein - to
27	8	1.0	855	2 T17460	disease resistance
28	8	1.0	863	2 A55173	cf-9 protein precu
29	8	1.0	910	2 B96770	hypothetical prote

receptor protein k  
protein kinase Xa2  
isoleucine--trna l  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
CAT8 protein - yea  
protein T7N9.24 [1  
ribosomal protein  
ribosomal protein  
hypothetical prote  
50S ribosomal prot  
cytochrome P450 (c  
phage hypothetical  
NADH dehydrogenase  
arsenical resistan  
ly-5-8 glycoprotei  
arsenical resistan  
probable membrane  
gene 61 protein -  
protein T2E6.14 [1  
probable membrane  
virulence regulato  
hypothetical prote  
hypothetical prote  
disease resistance  
transcription regu  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
maf protein TC0628  
alkyl hydroperoxid  
viral capsid prote  
1-acylglycerol-3-p  
1-acylglycerol-3-p  
hypothetical prote  
perA protein - Esc  
glucose inhibited  
glucose inhibited  
hypothetical prote  
hypothetical prote  
DNA-directed RNA p  
deoxyribose-phosph  
probable membrane  
ig light chain (I3  
hypothetical prote  
probable permease  
probable membrane  
hypothetical prote  
hypothetical prote  
rRNA (adenine-N6-)  
chlorophyll a/b-b1  
hypothetical prote  
conserved hypothet  
hypothetical prote  
probable adenine-s  
probable ABC trans  
protein F26P24.12  
ribose abc transpo  
hypothetical prote  
tail fiber [import  
hypothetical prote  
uncharacterized co  
transcription repr  
hypothetical prote  
stomatinal-like prot  
rRNA (adenine-N6-)  
probable S-adenosy  
probable enzyme [1

103	7	0.9	323	2	A85800	probable enzyme ye	176	7	0.9	566	2	S74633	high affinity sulf
104	7	0.9	323	2	AF0744	conserved hypoteth	177	7	0.9	568	2	S51275	DNA polymerase - p
105	7	0.9	327	1	DETWA	malate dehydrogena	178	7	0.9	572	2	E90183	hypothetical prote
106	7	0.9	335	2	H87247	probable cell divi	179	7	0.9	577	2	A83771	SNF2 helicase BH09
107	7	0.9	338	2	A64303	conserved hypoteth	180	7	0.9	589	2	T18239	transcription effe
108	7	0.9	341	1	K1BE36	thymidine kinase (	181	7	0.9	590	2	S63193	hypothetical prote
109	7	0.9	342	2	B53125	restriction enzyme	182	7	0.9	598	2	T25207	hypothetical prote
110	7	0.9	342	2	H81317	probable lipopolys	183	7	0.9	611	2	T28171	hypothetical prote
111	7	0.9	345	2	D84012	N-acetylglutamate	184	7	0.9	613	2	T15489	hypothetical prote
112	7	0.9	353	2	E98170	transcription regu	185	7	0.9	623	2	G81420	hypothetical prote
113	7	0.9	353	2	AH3116	transcription regu	186	7	0.9	624	2	T28423	ORF MSV261 leucine
114	7	0.9	360	2	D90657	hypothetical prote	187	7	0.9	625	2	T03837	surface lipoprotei
115	7	0.9	360	2	D85508	hypothetical prote	188	7	0.9	627	2	T03821	prolipoprotein p65
116	7	0.9	360	2	B86347	hypothetical prote	189	7	0.9	627	2	S74727	precorrin methylas
117	7	0.9	364	2	H87184	5'-phosphoribosyl-	190	7	0.9	630	2	T31798	hypothetical prote
118	7	0.9	364	2	E86456	unknown protein [l	191	7	0.9	630	2	T31017	hypothetical prote
119	7	0.9	369	2	S13721	Wnt-1 protein prec	192	7	0.9	638	2	T51383	receptor protein k
120	7	0.9	370	1	TVHUT1	transforming prote	193	7	0.9	641	2	C84726	probable receptor-
121	7	0.9	370	1	TVMSV1	transforming prote	194	7	0.9	654	2	B84690	hypothetical prote
122	7	0.9	370	1	TVMTV1	transforming prote	195	7	0.9	655	2	T30044	hypothetical prote
123	7	0.9	371	1	TVXLT1	transforming prote	196	7	0.9	657	2	S43415	histidine ammonia-
124	7	0.9	387	2	T04322	polysaccharonase	197	7	0.9	659	2	JC4910	heparitin-sulfate
125	7	0.9	387	2	JN0793	adaptive-response	198	7	0.9	662	2	G86428	F26G16.2 protein -
126	7	0.9	395	2	S47987	actin-related prot	199	7	0.9	662	2	G86210	hypothetical prote
127	7	0.9	402	2	A84581	probable disease r	200	7	0.9	662	2	T04856	hypothetical prote
128	7	0.9	403	2	T27825	hypothetical prote	201	7	0.9	665	2	AG1117	transketolase homo
129	7	0.9	410	2	A71481	probable poly A po	202	7	0.9	665	2	AI1477	transketolase homo
130	7	0.9	415	2	T13435	hypothetical prote	203	7	0.9	669	2	S64795	suppressor protein
131	7	0.9	417	2	T34459	hypothetical prote	204	7	0.9	674	2	T48261	hypothetical prote
132	7	0.9	417	2	F95953	probable glycosylt	205	7	0.9	676	2	A45515	dnak-type molecula
133	7	0.9	418	2	E91037	hypothetical prote	206	7	0.9	677	2	H86208	protein F20G5.26 l
134	7	0.9	418	2	B65013	hypothetical prote	207	7	0.9	689	2	AE1132	probable membrane
135	7	0.9	418	2	G85881	hypothetical prote	208	7	0.9	696	2	T42659	hypothetical prote
136	7	0.9	419	2	A90888	hypothetical prote	209	7	0.9	702	2	T21148	hypothetical prote
137	7	0.9	419	2	T49292	hypothetical prote	210	7	0.9	702	2	A86383	76.4K protein kina
138	7	0.9	420	1	H85729	hypothetical prote	211	7	0.9	707	2	T28418	ORF MSV257 leucine
139	7	0.9	420	1	DCECD	diaminopimelate de	212	7	0.9	725	2	T01268	leucine-rich repea
140	7	0.9	420	2	G91090	diaminopimelate de	213	7	0.9	728	2	D86278	hypothetical prote
141	7	0.9	420	2	B85936	diaminopimelate de	214	7	0.9	732	1	JC4655	acylaminoacyl-pept
142	7	0.9	421	2	T19315	hypothetical prote	215	7	0.9	732	1	JU0132	acylaminoacyl-pept
143	7	0.9	429	2	JC4986	site-specific DNA-	216	7	0.9	732	1	S07624	probable membrane
144	7	0.9	429	2	T28400	ORF MSV239 leucine	217	7	0.9	733	2	S56277	N-ethylmaleimide s
145	7	0.9	430	2	T04668	phosphoserine tran	218	7	0.9	739	2	T02996	hypothetical prote
146	7	0.9	438	2	H72241	fixc protein - The	219	7	0.9	740	2	B84741	probable disease r
147	7	0.9	440	2	F86667	UDP-N-acetylmuromo	220	7	0.9	741	2	T05250	probable disease r
148	7	0.9	440	2	G59102	hypothetical prote	221	7	0.9	744	2	C84527	probable receptor-
149	7	0.9	443	2	T28015	hypothetical prote	222	7	0.9	756	2	G86308	Similar to disease
150	7	0.9	447	2	JC5352	2-nitrotoluene dio	223	7	0.9	756	2	D75622	hypothetical prote
151	7	0.9	448	2	D81740	conserved hypoteth	224	7	0.9	757	2	T38124	hypothetical prote
152	7	0.9	449	2	G95004	competence factor	225	7	0.9	758	2	T16800	hypothetical prote
153	7	0.9	449	2	D97877	transport protein	226	7	0.9	761	2	S00703	photosystem I prot
154	7	0.9	454	2	AH0126	conserved hypoteth	227	7	0.9	764	1	H70414	conserved hypoteth
155	7	0.9	456	2	S14924	hypothetical prote	228	7	0.9	766	2	T01817	hypothetical prote
156	7	0.9	457	2	T22173	hypothetical prote	229	7	0.9	771	2	T02565	disease resistance
157	7	0.9	462	2	A40552	bindin fertilizati	230	7	0.9	773	2	D90099	hypothetical prote
158	7	0.9	463	2	T10065	phosphoadenylyl-su	231	7	0.9	791	2	T41573	hypothetical prote
159	7	0.9	477	2	T41935	capsid protein - h	232	7	0.9	799	2	JH0797	castor protein - f
160	7	0.9	480	2	T00971	probable disease r	233	7	0.9	818	2	F96586	hypothetical prote
161	7	0.9	494	2	E71055	hypothetical prote	234	7	0.9	830	2	T34472	hypothetical prote
162	7	0.9	509	2	D69832	probable Rieske [2	235	7	0.9	830	2	S50810	probable membrane
163	7	0.9	511	2	JC5980	L-2,4-diaminobuty	236	7	0.9	839	2	T12827	hypothetical prote
164	7	0.9	512	1	O4HU6	aryl hydrocarbon (	237	7	0.9	845	2	T07039	hcr9-0 protein - t
165	7	0.9	512	2	S21761	aryl hydrocarbon (	238	7	0.9	845	2	G86708	aconitate hydratase
166	7	0.9	512	2	G86459	Hypothetical 55.6	239	7	0.9	851	2	AD1427	internalin, probab
167	7	0.9	513	1	I58311	HMG-box containing	240	7	0.9	852	2	A72343	hypothetical prote
168	7	0.9	531	2	T04463	hypothetical prote	241	7	0.9	858	2	T00258	hypothetical prote
169	7	0.9	540	2	T12704	leucine-rich prote	242	7	0.9	871	2	D96557	probable protein k
170	7	0.9	545	2	T02578	hypothetical prote	243	7	0.9	875	2	H96557	probable protein k
171	7	0.9	549	1	S39533	phosphoprotein pho	244	7	0.9	880	2	S60137	beta-N-acetylhexos
172	7	0.9	549	2	T41744	hypothetical prote	245	7	0.9	883	2	E96557	probable protein k
173	7	0.9	555	2	D95139	DNA repair protein	246	7	0.9	884	2	T02731	serine/threonine-s
174	7	0.9	555	2	C98007	DNA repair and gen	247	7	0.9	885	2	B69783	transporter homolo
175	7	0.9	562	2	T34319	hypothetical prote	248	7	0.9	890	2	T00800	disease resistance

249	7	0.9	892	2	T01899	disease resistance	322	7	0.9	4307	2	T20721	hypothetical prote
250	7	0.9	893	2	H96651	protein T3P18.19 [	323	7	0.9	4367	1	B54802	myelin heavy chain
251	7	0.9	904	2	T46170	disease resistance	324	7	0.9	4391	2	A38096	perlecan precursor
252	7	0.9	905	2	T00475	probable disease r	325	7	0.9	4717	2	T41581	hypothetical coile
253	7	0.9	907	2	A86460	99.9k hypothetical	326	6	0.7	26	2	B24743	prolactin, 24K - M
254	7	0.9	949	2	D97781	hypothetical prote	327	6	0.7	39	2	I55273	DNA-directed DNA p
255	7	0.9	949	2	F90086	hypothetical prote	328	6	0.7	45	2	S77773	probable thymidine
256	7	0.9	950	2	E70203	chromosomal region	329	6	0.7	49	2	S33347	complement C4 prot
257	7	0.9	951	2	A96770	hypothetical prote	330	6	0.7	49	2	D95054	hypothetical prote
258	7	0.9	960	2	T50383	homolog to yeast c	331	6	0.7	49	2	A37405	androgen-binding p
259	7	0.9	964	2	T49038	hypothetical prote	332	6	0.7	56	2	G90278	hypothetical prote
260	7	0.9	976	2	B84659	probable receptor-	333	6	0.7	60	2	A85598	hypothetical prote
261	7	0.9	988	2	T45717	receptor-kinase l1	334	6	0.7	61	2	PS0404	muscle segment hom
262	7	0.9	992	2	T05335	hypothetical prote	335	6	0.7	61	2	PS0404	muscle segment hom
263	7	0.9	999	1	S27756	receptor-like prot	336	6	0.7	64	2	T15163	hypothetical prote
264	7	0.9	1002	2	T46033	receptor protein k	337	6	0.7	64	2	E97815	hypothetical prote
265	7	0.9	1007	2	C84668	probable receptor-	338	6	0.7	65	2	E70644	probable ribosomal
266	7	0.9	1019	2	C96519	probable disease r	339	6	0.7	65	2	S08645	core antigen (clon
267	7	0.9	1021	2	A86421	Receptor-like seri	340	6	0.7	65	2	S08646	core antigen (clon
268	7	0.9	1025	2	S50293	probable membrane	341	6	0.7	65	2	S08647	core antigen (clon
269	7	0.9	1025	2	T42626	secreted leucine-r	342	6	0.7	65	2	S08078	core antigen (clon
270	7	0.9	1027	2	B85089	receptor protein k	343	6	0.7	65	2	AC3139	hypothetical prote
271	7	0.9	1051	2	T13174	gp150 protein - fr	344	6	0.7	67	2	S33341	complement C4 prot
272	7	0.9	1066	2	T15864	hypothetical prote	345	6	0.7	67	2	S33342	complement C4 prot
273	7	0.9	1078	2	D87647	hypothetical prote	346	6	0.7	67	2	AH0984	conserved hypothet
274	7	0.9	1079	2	C96772	probable receptor	347	6	0.7	68	2	T37135	hypothetical prote
275	7	0.9	1121	2	S54504	hypothetical prote	348	6	0.7	68	2	AF0066	hypothetical prote
276	7	0.9	1122	2	A97814	transcription-repa	349	6	0.7	69	2	B89428	protein T08D2.6 [1
277	7	0.9	1123	2	D96756	receptor-like prot	350	6	0.7	71	2	T45384	ribosomal protein
278	7	0.9	1125	2	T19193	hypothetical prote	351	6	0.7	71	2	T28271	ORF MSV109 hypothe
279	7	0.9	1166	2	F96598	protein F20N2.4 [1	352	6	0.7	72	2	G64130	hypothetical prote
280	7	0.9	1184	2	H71436	hypothetical prote	353	6	0.7	73	2	H82609	hypothetical prote
281	7	0.9	1189	2	A54817	AtPase ScII, chrom	354	6	0.7	74	2	AB2071	hypothetical prote
282	7	0.9	1192	2	T48499	receptor-like prot	355	6	0.7	74	2	AH3307	transcription regu
283	7	0.9	1196	2	T09356	brassinosteroid-in	356	6	0.7	75	2	H72236	conserved hypothet
284	7	0.9	1203	2	B55094	chromosomal protei	357	6	0.7	75	2	T08508	trbk protein - Ent
285	7	0.9	1209	2	A49440	chromosome disjunc	358	6	0.7	78	2	D91160	hypothetical prote
286	7	0.9	1237	2	AC1583	internalin protein	359	6	0.7	78	2	E65136	hypothetical 8.7 k
287	7	0.9	1238	2	AH0038	probable exported	360	6	0.7	78	2	C86006	hypothetical prote
288	7	0.9	1256	2	S60461	gene flightless-I	361	6	0.7	79	2	G34510	homeotic protein H
289	7	0.9	1268	2	AB0204	conserved hypothet	362	6	0.7	79	2	S72305	conserved hypothet
290	7	0.9	1274	2	T02636	D1 protein homolog	363	6	0.7	79	2	G82709	hypothetical prote
291	7	0.9	1291	1	A28334	protein-tyrosine-p	364	6	0.7	83	2	AC3605	hypothetical prote
292	7	0.9	1301	2	D85188	disease resistance	365	6	0.7	84	2	E86926	hypothetical prote
293	7	0.9	1304	2	G85188	disease resistance	366	6	0.7	84	2	C97995	degenerate transpo
294	7	0.9	1314	1	TNBYR6	transcription regu	367	6	0.7	85	2	S56097	alpha-glucosidase
295	7	0.9	1317	2	B85189	disease resistance	368	6	0.7	85	2	T15419	FMRFamide-like pep
296	7	0.9	1378	1	A48751	protein-tyrosine k	369	6	0.7	86	2	D89862	hypothetical prote
297	7	0.9	1379	2	S78371	DNA-directed RNA p	370	6	0.7	88	2	AB6259	protein F5011.20 [
298	7	0.9	1400	1	I38185	protein-tyrosine k	371	6	0.7	88	2	C97251	probable membrane
299	7	0.9	1422	2	B71437	probable resistanc	372	6	0.7	89	2	D83052	30S ribosomal prot
300	7	0.9	1436	2	S67655	probable resistanc	373	6	0.7	89	2	G97491	30S ribosomal prot
301	7	0.9	1448	2	F83237	probable ATP-depen	374	6	0.7	89	2	AF2709	hypothetical prote
302	7	0.9	1454	2	S53398	hypothetical prote	375	6	0.7	92	2	B42219	hypothetical prote
303	7	0.9	1495	2	T31434	densin-180 - rat	376	6	0.7	92	2	D64039	hypothetical prote
304	7	0.9	1531	2	T42218	slit-1 protein hom	377	6	0.7	92	2	AB2021	hypothetical prote
305	7	0.9	1592	2	S63208	hypothetical prote	378	6	0.7	94	2	S17449	probable ligand-bi
306	7	0.9	1724	2	T13942	UNC-13-B protein -	379	6	0.7	97	2	H84167	hypothetical prote
307	7	0.9	1726	2	T30810	chromatin structur	380	6	0.7	98	2	S04755	NADH dehydrogenase
308	7	0.9	1813	2	T30564	resistance protein	381	6	0.7	98	2	S60229	gibberellin-regula
309	7	0.9	1968	1	PN0093	genome polypotein	382	6	0.7	98	4	B36445	gibberellin-regula
310	7	0.9	2018	2	T34274	hypothetical prote	383	6	0.7	99	2	S60230	gibberellin-regula
311	7	0.9	2026	1	OYBY	adenylate cyclase	384	6	0.7	99	2	S60231	hypothetical prote
312	7	0.9	2049	2	T47587	TATA box binding p	385	6	0.7	99	2	B83392	hypothetical prote
313	7	0.9	2057	2	F90109	splicing factor Pr	386	6	0.7	99	2	T43600	probable transposa
314	7	0.9	2062	2	G96602	probable receptor	387	6	0.7	99	2	A82651	hypothetical prote
315	7	0.9	2198	2	T20371	hypothetical prote	388	6	0.7	100	2	T08248	TATA-binding trans
316	7	0.9	2204	1	REN2NV	genome polypotein	389	6	0.7	100	2	A70875	probable PE protei
317	7	0.9	2210	1	RRXPVY	genome polypotein	390	6	0.7	101	2	T48791	hypothetical prote
318	7	0.9	2467	2	D71437	probable resistanc	391	6	0.7	101	2	A99974	hypothetical prote
319	7	0.9	3036	2	T18995	hypothetical prote	392	6	0.7	101	2	AC0099	probable lipoprote
320	7	0.9	3707	2	S18252	heparan sulfate pr	393	6	0.7	101	2	F90969	probable minor tai
321	7	0.9	4199	2	S76412	hypothetical prote	394	6	0.7	102	2	F69475	conserved hypothet

395	6	0.7	102	2	S48012	Ip7 protein - phag	468	6	0.7	133	2	T42978	hypothetical prote
396	6	0.7	102	2	Jc1150	hypothetical prote	469	6	0.7	133	2	B90260	conserved hypothet
397	6	0.7	103	2	T31207	hypothetical prote	470	6	0.7	134	2	G84366	hypothetical prote
398	6	0.7	105	2	H97704	hypothetical prote	471	6	0.7	135	2	164052	probable DNA-bindi
399	6	0.7	106	2	P00871	cell fusion protei	472	6	0.7	135	2	B81097	transcription regu
400	6	0.7	106	2	P00873	cell fusion protei	473	6	0.7	136	2	JQ1359	C2 protein - Misca
401	6	0.7	106	2	P00869	cell fusion protei	474	6	0.7	136	2	T33895	hypothetical prote
402	6	0.7	106	2	G82729	hypothetical prote	475	6	0.7	136	2	G82928	ATP synthase epsil
403	6	0.7	106	2	T10097	nifX protein - Met	476	6	0.7	137	2	180176	class I histocompa
404	6	0.7	107	2	S52508	probable membrane	477	6	0.7	137	2	185718	probable tail comp
405	6	0.7	107	2	G83348	hypothetical prote	478	6	0.7	137	2	D97544	transcription regu
406	6	0.7	109	2	S66530	endothelin convert	479	6	0.7	137	2	AE2763	transcription regu
407	6	0.7	111	2	G70521	hypothetical prote	480	6	0.7	138	2	S73943	MG076 homolog G07-
408	6	0.7	112	2	E60767	retrovirus-related	481	6	0.7	139	2	H90032	hypothetical prote
409	6	0.7	112	2	B86441	hypothetical prote	482	6	0.7	139	2	A38612	insulin-like growt
410	6	0.7	113	2	A81093	hypothetical prote	483	6	0.7	140	2	AC1280	hypothetical prote
411	6	0.7	113	2	E81848	hypothetical prote	484	6	0.7	140	2	AC1643	hypothetical prote
412	6	0.7	113	2	F64001	hypothetical prote	485	6	0.7	140	2	AH1454	protein from Bacte
413	6	0.7	113	2	D82250	phosphorelay prote	486	6	0.7	140	2	A11090	a protein from Bac
414	6	0.7	114	2	JC5482	hypothetical 13.2K	487	6	0.7	141	2	T12115	NADH dehydrogenase
415	6	0.7	114	2	B83709	hypothetical prote	488	6	0.7	141	2	AC1994	hypothetical prote
416	6	0.7	114	2	B97036	probable metal-bin	489	6	0.7	142	2	B82308	ribosomal protein
417	6	0.7	115	2	T17350	NADH dehydrogenase	490	6	0.7	142	2	G75171	isu ribosomal prot
418	6	0.7	117	2	B97835	hypothetical prote	491	6	0.7	142	2	T28987	hypothetical prote
419	6	0.7	117	2	S77724	hypothetical prote	492	6	0.7	142	2	D72340	hypothetical prote
420	6	0.7	118	1	B42959	14K hypothetical p	493	6	0.7	143	1	FEYTA	ferredoxin [3Fe-4S
421	6	0.7	118	2	D97498	hypothetical prote	494	6	0.7	143	1	A60159	interleukin-3 prec
422	6	0.7	118	2	E91045	hypothetical prote	495	6	0.7	143	2	A31574	adrenodoxin homolo
423	6	0.7	118	2	S73487	RNaseP C5 chain -	496	6	0.7	143	2	E72349	conserved hypothet
424	6	0.7	118	2	H85889	hypothetical prote	497	6	0.7	143	2	AG1127	hypothetical prote
425	6	0.7	118	2	AF0860	hypothetical prote	498	6	0.7	143	2	T49457	hypothetical prote
426	6	0.7	119	2	JC4546	hypothetical 13.7k	499	6	0.7	144	2	B90834	minor tail protein
427	6	0.7	119	2	B72526	hypothetical prote	500	6	0.7	144	2	D85691	probable tail comp
428	6	0.7	119	2	G90324	conserved hypothet	501	6	0.7	145	2	E84461	hypothetical prote
429	6	0.7	120	1	S66530	crossover junction	502	6	0.7	145	2	C90107	hypothetical prote
430	6	0.7	120	2	G90831	hypothetical prote	503	6	0.7	145	2	AE3531	zinc uptake regula
431	6	0.7	120	2	F85688	hypothetical prote	504	6	0.7	146	2	E97226	aspartate carbamoy
432	6	0.7	120	2	F82153	hypothetical prote	505	6	0.7	147	2	A72248	ribosomal protein
433	6	0.7	120	2	T05659	hypothetical prote	506	6	0.7	147	2	E83677	hypothetical prote
434	6	0.7	121	2	H83687	hypothetical prote	507	6	0.7	147	2	S76050	hypothetical prote
435	6	0.7	122	2	T03087	hypothetical prote	508	6	0.7	147	2	C71802	hypothetical prote
436	6	0.7	122	2	G82606	conjugal transfer	509	6	0.7	147	2	F64715	hypothetical prote
437	6	0.7	122	2	G64558	hypothetical prote	510	6	0.7	147	2	H75319	hypothetical prote
438	6	0.7	122	2	A71950	hypothetical prote	511	6	0.7	147	2	S74460	hypothetical prote
439	6	0.7	122	2	D90425	conserved hypothet	512	6	0.7	147	2	F81193	hypothetical prote
440	6	0.7	123	2	S67584	probable membrane	513	6	0.7	148	2	G95034	ribosomal protein
441	6	0.7	123	2	T09288	probable tail comp	514	6	0.7	148	2	G97905	50S ribosomal prot
442	6	0.7	123	2	AH1998	hypothetical prote	515	6	0.7	148	2	H64318	transcription regu
443	6	0.7	124	2	D64219	ribosomal protein	516	6	0.7	148	2	T32810	hypothetical prote
444	6	0.7	124	2	A84313	glycine decarboxyl	517	6	0.7	148	2	E69032	formate hydrogenly
445	6	0.7	124	2	S70043	hypothetical prote	518	6	0.7	149	2	E69967	transcription regu
446	6	0.7	124	2	AF2403	hypothetical prote	519	6	0.7	149	2	F64326	hypothetical prote
447	6	0.7	124	2	G86258	protein F501.16 l	520	6	0.7	149	2	B86475	hypothetical prote
448	6	0.7	125	2	A72566	hypothetical prote	521	6	0.7	150	2	G83840	ferric uptake regu
449	6	0.7	125	2	S76216	hypothetical prote	522	6	0.7	150	2	S69526	hypothetical prote
450	6	0.7	125	2	T25872	hypothetical prote	523	6	0.7	150	2	T08734	hypothetical prote
451	6	0.7	126	2	T18476	hypothetical prote	524	6	0.7	151	2	JN0521	toxin co-regulated
452	6	0.7	126	2	AT2716	proteinase VII pre	525	6	0.7	152	1	A24427	interleukin-3 prec
453	6	0.7	126	2	F83556	probable ring-clea	526	6	0.7	152	1	S42720	interleukin-3 prec
454	6	0.7	126	2	F89989	hypothetical prote	527	6	0.7	152	1	B24427	interleukin-3 prec
455	6	0.7	127	2	PC2263	cytochrome P450 pr	528	6	0.7	152	2	AD0964	probable exported
456	6	0.7	127	2	PC2261	cytochrome P450 pr	529	6	0.7	152	2	AB1139	transcription regu
457	6	0.7	127	2	PC2262	cytochrome P450 pr	530	6	0.7	152	2	AB11496	weakly transcripti
458	6	0.7	128	2	G75308	hypothetical prote	531	6	0.7	152	2	AG1774	hypothetical prote
459	6	0.7	128	2	S75603	hypothetical prote	532	6	0.7	153	2	D84706	hypothetical prote
460	6	0.7	129	2	B64510	hypothetical prote	533	6	0.7	154	1	A70177	cytidine deaminase
461	6	0.7	129	2	A90768	probable minor tai	534	6	0.7	154	2	AH1585	hypothetical prote
462	6	0.7	129	2	A96787	protein F10A5.6 [i	535	6	0.7	154	2	A11231	hypothetical prote
463	6	0.7	129	2	D82627	hypothetical prote	536	6	0.7	154	2	F86437	protein P28K20.12
464	6	0.7	130	2	S24340	H+-transporting AT	537	6	0.7	155	2	H69054	hypothetical prote
465	6	0.7	130	2	T16101	hypothetical prote	538	6	0.7	155	2	AF1195	hypothetical prote
466	6	0.7	132	1	A69311	conserved hypothet	539	6	0.7	155	2	AC2032	heat shock protein
467	6	0.7	132	2	D81381	transcription term	540	6	0.7	156	2	C83390	cytanease lyase PA20

541	6	0.7	156	2	T26634	hypothetical prote	614	6	0.7	179	2	E75189	hypothetical prote
542	6	0.7	157	1	ZNPRT9	recombination endo	615	6	0.7	179	2	T48416	hypothetical prote
543	6	0.7	157	2	AG1712	PTS system, fructo	616	6	0.7	179	2	D90167	conserved hypothet
544	6	0.7	157	2	A11341	PTS system, fructo	617	6	0.7	180	1	GCRTDU	glucagon precursor
545	6	0.7	158	2	A4948	ImbB protein - St	618	6	0.7	181	2	S53178	core antigen - hep
546	6	0.7	158	2	S64321	hypothetical prote	619	6	0.7	181	2	E36794	hypothetical prote
547	6	0.7	159	2	D97338	biotin carboxyl c	620	6	0.7	181	2	A69540	conserved hypothet
548	6	0.7	159	2	AH2118	bacterioferritin c	621	6	0.7	182	2	T35124	hypothetical prote
549	6	0.7	159	2	T48956	hypothetical prote	622	6	0.7	182	2	T20185	hypothetical prote
550	6	0.7	159	2	A72709	hypothetical prote	623	6	0.7	182	2	F71072	hypothetical prote
551	6	0.7	159	2	S61040	probable membrane	624	6	0.7	183	1	NKVL42	core antigen - hep
552	6	0.7	160	2	G71122	hypothetical prote	625	6	0.7	183	1	NKVLCP	core antigen - hep
553	6	0.7	160	2	A35261	Alginatase regulator	626	6	0.7	183	2	S54818	ribulose-bisphosph
554	6	0.7	161	2	S30698	cyax protein - Esc	627	6	0.7	183	2	S53214	core antigen - hep
555	6	0.7	161	2	B91221	hypothetical prote	628	6	0.7	183	2	S53221	core antigen - hep
556	6	0.7	161	2	D86067	hypothetical prote	629	6	0.7	183	2	S53232	core antigen - hep
557	6	0.7	161	2	S64869	probable membrane	630	6	0.7	183	2	S53260	core antigen - hep
558	6	0.7	161	2	H91235	2-demethylmenaquin	631	6	0.7	183	2	S53267	core antigen - hep
559	6	0.7	161	2	S40872	2-demethylmenaquin	632	6	0.7	183	2	S53270	core antigen - hep
560	6	0.7	161	2	T47390	hypothetical prote	633	6	0.7	183	2	S53286	core antigen - hep
561	6	0.7	161	2	H86082	menaquinone biosyn	634	6	0.7	183	2	S43490	core antigen - hep
562	6	0.7	161	2	AD0939	menaquinone biosyn	635	6	0.7	183	2	S53129	core antigen - hep
563	6	0.7	162	2	A64972	probable acetyl tr	636	6	0.7	183	2	S53137	core antigen - hep
564	6	0.7	162	2	C71968	Shikimate kinase I	637	6	0.7	183	2	S53140	core antigen - hep
565	6	0.7	162	2	E64539	shikimic acid kina	638	6	0.7	183	2	S53143	core antigen - hep
566	6	0.7	162	2	T16102	hypothetical prote	639	6	0.7	183	2	S53152	core antigen - hep
567	6	0.7	162	2	G90986	hypothetical prote	640	6	0.7	183	2	S53155	core antigen - hep
568	6	0.7	162	2	B85832	probable transfera	641	6	0.7	183	2	S53169	core antigen - hep
569	6	0.7	162	2	AG0769	probable transfera	642	6	0.7	183	2	S53172	core antigen - hep
570	6	0.7	163	2	S33345	probable acetyltra	643	6	0.7	183	2	S53175	core antigen - hep
571	6	0.7	163	2	S33340	complement C4 prot	644	6	0.7	183	2	S53184	core antigen - hep
572	6	0.7	163	2	S33348	complement C4 prot	645	6	0.7	183	2	S53189	core antigen - hep
573	6	0.7	163	2	S33418	complement C4 prot	646	6	0.7	183	2	S53181	core antigen - hep
574	6	0.7	163	2	S33343	complement C4 prot	647	6	0.7	183	2	S22318	core antigen - hep
575	6	0.7	163	2	S33344	complement C4 prot	648	6	0.7	183	2	D84134	hypothetical prote
576	6	0.7	163	2	F71b35	hypothetical prote	649	6	0.7	183	2	T26637	hypothetical prote
577	6	0.7	164	2	S01280	hypothetical prote	650	6	0.7	184	2	S53146	core antigen - hep
578	6	0.7	164	2	E75100	hypothetical prote	651	6	0.7	184	2	S61304	transcription acti
579	6	0.7	166	2	G98118	hypothetical prote	652	6	0.7	184	2	AG2344	hypothetical prote
580	6	0.7	167	2	D70385	hypothetical prote	653	6	0.7	184	2	C75021	transcription regu
581	6	0.7	167	2	G95341	protein [imported	654	6	0.7	184	2	A11589	hypothetical prote
582	6	0.7	168	2	H85071	hypothetical prote	655	6	0.7	185	2	S53288	core antigen - hep
583	6	0.7	168	2	D82703	7,8-dihydro-8-oxog	656	6	0.7	185	2	C82321	conserved hypothet
584	6	0.7	169	2	B84511	hypothetical prote	657	6	0.7	185	2	B72212	translation elonga
585	6	0.7	169	2	A72466	hypothetical prote	658	6	0.7	186	2	JC5354	2-nitrotoluene dio
586	6	0.7	170	2	I61888	hypothetical prote	659	6	0.7	186	2	T30918	hypothetical prote
587	6	0.7	170	2	I61887	complement C4 - ch	660	6	0.7	186	2	AC1650	hypothetical prote
588	6	0.7	170	2	I84442	complement C4 - or	661	6	0.7	187	2	S58668	positive regulator
589	6	0.7	170	2	I61890	complement C4 - ch	662	6	0.7	187	2	D88499	protein K04G7.2 [i
590	6	0.7	170	2	I37024	complement C4 - go	663	6	0.7	188	2	S75574	transposase sll080
591	6	0.7	170	2	B71510	probable methyltra	664	6	0.7	188	2	G75326	probable 16S rRNA
592	6	0.7	171	2	A10943	probable membrane	665	6	0.7	189	2	JC7107	development relate
593	6	0.7	172	2	B83596	hypothetical prote	666	6	0.7	189	2	T18480	hypothetical prote
594	6	0.7	173	2	C84684	hypothetical prote	667	6	0.7	190	2	A82343	conserved hypothet
595	6	0.7	173	2	I51269	lutropin receptor	668	6	0.7	190	2	AG1222	hypothetical prote
596	6	0.7	174	2	B54853	amastigote surface	669	6	0.7	191	1	B64129	dethiobiotin synth
597	6	0.7	174	2	C54853	amastigote surface	670	6	0.7	191	2	B40628	resolvase homolog
598	6	0.7	174	2	A54853	amastigote surface	671	6	0.7	191	2	C81279	probable molybdopt
599	6	0.7	174	2	F89813	conserved hypothet	672	6	0.7	193	2	T06124	hypothetical prote
600	6	0.7	174	2	AH3296	deoxyuridine 5'-tr	673	6	0.7	193	2	S22881	gonadal protein gd
601	6	0.7	175	2	T50849	hypothetical prote	674	6	0.7	193	2	G88039	protein F53A10.1 [
602	6	0.7	176	2	A56652	interleukin-7 prec	675	6	0.7	193	2	A83280	hypothetical prote
603	6	0.7	176	2	JC4035	interleukin 7 prot	676	6	0.7	193	2	S26597	hypothetical prote
604	6	0.7	176	2	S77468	bacterioferritin I	677	6	0.7	194	2	T16556	hypothetical prote
605	6	0.7	176	2	E90160	conserved hypothet	678	6	0.7	194	2	E75083	hypothetical prote
606	6	0.7	176	2	T08254	probable transposa	679	6	0.7	195	1	NKVLH3	core antigen - hep
607	6	0.7	176	2	T23475	hypothetical prote	680	6	0.7	195	2	T49012	calmodulin-like pr
608	6	0.7	176	2	AD0685	hypothetical prote	681	6	0.7	195	2	B69298	orotate phosphorib
609	6	0.7	177	2	S74636	hypothetical prote	682	6	0.7	195	2	T29398	hypothetical prote
610	6	0.7	177	2	S18755	hypothetical prote	683	6	0.7	196	2	F97096	ribosomal protein
611	6	0.7	177	2	T49816	hypothetical prote	684	6	0.7	196	2	S73870	Mg208 homolog H10
612	6	0.7	178	2	D72281	hypothetical prote	685	6	0.7	196	2	I64222	conserved hypothet
613	6	0.7	179	2	F81087	hypothetical prote	686	6	0.7	196	2	AC0488	two-component syst

687	6	0.7	197	2	H87446	potassium-transpor	760	6	0.7	212	2	S53274	e antigen precurs
688	6	0.7	197	2	S61047	hypothetical prote	761	6	0.7	212	2	S53279	e antigen precurs
689	6	0.7	197	2	S74851	hypothetical prote	762	6	0.7	212	2	S53281	e antigen precurs
690	6	0.7	197	2	S73759	MG286 homolog fl1	763	6	0.7	212	2	S20746	e antigen precurs
691	6	0.7	197	2	C69777	hypothetical prote	764	6	0.7	212	2	S25651	e antigen precurs
692	6	0.7	197	2	B72471	hypothetical prote	765	6	0.7	212	2	S32204	e antigen precurs
693	6	0.7	197	2	F70210	hypothetical prote	766	6	0.7	212	2	S20750	e antigen precurs
694	6	0.7	198	2	T06261	probable imidazole	767	6	0.7	212	2	S53159	e antigen precurs
695	6	0.7	198	2	A97404	hypothetical prote	768	6	0.7	212	2	S53163	e antigen precurs
696	6	0.7	198	2	A12621	conserved hypothet	769	6	0.7	212	2	S53198	e antigen precurs
697	6	0.7	199	2	D81985	probable pilus ass	770	6	0.7	212	2	S53200	e antigen precurs
698	6	0.7	199	2	T32267	hypothetical prote	771	6	0.7	212	2	S53202	e antigen precurs
699	6	0.7	200	2	B83539	hypothetical prote	772	6	0.7	212	2	JQ2227	e antigen precurs
700	6	0.7	200	2	AB0620	probable bacteriop	773	6	0.7	212	2	S53204	e antigen precurs
701	6	0.7	200	2	T18479	hypothetical prote	774	6	0.7	212	2	S53227	e antigen precurs
702	6	0.7	201	2	T15145	hypothetical prote	775	6	0.7	212	2	S67504	e antigen precurs
703	6	0.7	202	2	A82717	50S ribosomal prot	776	6	0.7	212	2	F65023	gcvR protein - Esc
704	6	0.7	202	2	S76283	hypothetical prote	777	6	0.7	212	2	E91046	transcription regu
705	6	0.7	202	2	T30927	hypothetical prote	778	6	0.7	212	2	A85891	transcriptional regu
706	6	0.7	202	2	T06621	hypothetical prote	779	6	0.7	212	2	T44591	hypothetical prote
707	6	0.7	202	2	T49134	hypothetical prote	780	6	0.7	212	2	C90905	hypothetical prote
708	6	0.7	204	2	D69349	conserved hypothet	781	6	0.7	212	2	H98237	biolod protein (AF31
709	6	0.7	204	2	T15274	hypothetical prote	782	6	0.7	212	2	AC3048	dethiolotin synth
710	6	0.7	205	2	S29251	hypothetical prote	783	6	0.7	212	2	AH3606	dethiolotin synth
711	6	0.7	205	2	G70338	conserved hypothet	784	6	0.7	213	1	TVHUC4	transforming prote
712	6	0.7	205	2	G90141	conserved hypothet	785	6	0.7	213	1	AK0581	conserved hypothet
713	6	0.7	205	2	D83436	probable sugar ald	786	6	0.7	214	1	NKVLA3	e antigen precurs
714	6	0.7	205	2	H64001	hypothetical prote	787	6	0.7	214	1	NKVLA6	e antigen precurs
715	6	0.7	205	2	AB1903	hypothetical prote	788	6	0.7	214	1	NKVLS	e antigen precurs
716	6	0.7	206	1	E70197	conserved hypothet	789	6	0.7	214	2	T20294	hypothetical prote
717	6	0.7	206	2	T07697	ribosomal protein	790	6	0.7	214	2	T09535	dnak-type molecula
718	6	0.7	206	2	T04517	hypothetical prote	791	6	0.7	214	2	S01405	core antigen - hep
719	6	0.7	206	2	S73084	hypothetical prote	792	6	0.7	214	2	S47409	e antigen precurs
720	6	0.7	207	2	AI0002	secreted thiol,dis	793	6	0.7	214	2	S33686	e antigen precurs
721	6	0.7	207	2	AB0951	thiol,disulfide in	794	6	0.7	214	2	T42530	hypothetical prote
722	6	0.7	207	2	C97221	hypothetical prote	795	6	0.7	214	2	A97540	hypothetical prote
723	6	0.7	207	2	B81418	hypothetical prote	796	6	0.7	214	2	AC2759	cytochrome b561 At
724	6	0.7	207	2	A69941	capsular polysacch	797	6	0.7	215	2	T25369	hypothetical prote
725	6	0.7	208	2	G69901	acyl-carrier prote	798	6	0.7	215	2	H97154	sporulation factor
726	6	0.7	208	2	D83241	probable nuclease	799	6	0.7	216	2	AD0003	conserved hypothet
727	6	0.7	209	2	AC0192	probable GerR-fam1	800	6	0.7	216	2	S74711	hypothetical prote
728	6	0.7	209	2	S54209	zEM1 protein - mal	801	6	0.7	216	2	C75102	hypothetical prote
729	6	0.7	209	2	B90681	probable transcrip	802	6	0.7	216	2	B85712	unknown protein en
730	6	0.7	209	2	F85531	probable transcrip	803	6	0.7	216	2	F90393	phosphoglycerate m
731	6	0.7	210	2	AG0950	ATP/GTP-binding pr	804	6	0.7	216	2	D96946	transcription regu
732	6	0.7	210	2	S76008	ATP phosphoribosyl	805	6	0.7	217	2	T11382	cytochrome-c oxida
733	6	0.7	210	2	F81583	conserved hypothet	806	6	0.7	217	2	H72490	hypothetical prote
734	6	0.7	210	2	B71527	probable o-sialogl	807	6	0.7	217	2	A70166	hypothetical prote
735	6	0.7	210	2	T49170	hypothetical prote	808	6	0.7	218	2	A46048	glutathione transf
736	6	0.7	210	2	AD0547	RhtC-like transpor	809	6	0.7	218	2	H71233	hypothetical prote
737	6	0.7	211	1	NKVLA1	e antigen precurs	810	6	0.7	218	2	H71866	hypothetical prote
738	6	0.7	211	2	E70428	recombination prot	811	6	0.7	220	2	T23817	hypothetical prote
739	6	0.7	211	2	A64490	precorrin-6V methy	812	6	0.7	220	2	AD0225	probable amino-aci
740	6	0.7	212	1	NKVLAH	e antigen precurs	813	6	0.7	220	2	F86323	protein F14D16.13
741	6	0.7	212	1	NKVLBH	e antigen precurs	814	6	0.7	221	2	AI0360	conserved hypothet
742	6	0.7	212	1	NKVLA4	e antigen precurs	815	6	0.7	221	2	T07079	leucine-rich repea
743	6	0.7	212	1	NKVLAJ1	e antigen precurs	816	6	0.7	222	2	AE1553	B. subtilis yjbm p
744	6	0.7	212	1	NKVLAJ2	e antigen precurs	817	6	0.7	222	2	AG1195	B. subtilis yjbm p
745	6	0.7	212	2	A64045	KHG-KDPG bifunctio	818	6	0.7	222	2	H70192	endonuclease III (
746	6	0.7	212	2	S53211	e antigen precurs	819	6	0.7	222	2	AF0395	probable amino aci
747	6	0.7	212	2	S53216	e antigen precurs	820	6	0.7	222	2	T27939	hypothetical prote
748	6	0.7	212	2	S53223	e antigen precurs	821	6	0.7	223	2	H64759	membrane protein y
749	6	0.7	212	2	S53225	e antigen precurs	822	6	0.7	223	2	F90676	probable cytochrom
750	6	0.7	212	2	S53229	e antigen precurs	823	6	0.7	223	2	C69215	conserved hypothet
751	6	0.7	212	2	S53236	e antigen precurs	824	6	0.7	223	2	D84007	hypothetical prote
752	6	0.7	212	2	S53238	e antigen precurs	825	6	0.7	223	2	A85527	probable cytochrom
753	6	0.7	212	2	S53240	e antigen precurs	826	6	0.7	224	2	C96903	orotate phosphorib
754	6	0.7	212	2	S53242	e antigen precurs	827	6	0.7	224	2	A75170	ribonuclease h11 (
755	6	0.7	212	2	S53251	e antigen precurs	828	6	0.7	224	2	B84131	uracil-DNA glycosy
756	6	0.7	212	2	S53253	e antigen precurs	829	6	0.7	224	2	S48671	proliferin - human
757	6	0.7	212	2	S53255	e antigen precurs	830	6	0.7	224	2	A05086	proliferin 1 precu
758	6	0.7	212	2	S53257	e antigen precurs	831	6	0.7	224	2	AG2158	hypothetical prote
759	6	0.7	212	2	S53272	e antigen precurs	832	6	0.7	224	2	G70119	hypothetical prote



833	6	0.7	224	2	E97430	hypothetical prote	906	6	0.7	243	2	T25190	hypothetical prote
834	6	0.7	224	2	AF2648	telchoic acid bios	907	6	0.7	243	2	T47410	leucine-rich repea
835	6	0.7	225	2	S73560	phosphate transpor	908	6	0.7	244	2	A22732	proliferin-related
836	6	0.7	225	2	T38196	probable proteasom	909	6	0.7	244	2	S72896	hypothetical prote
837	6	0.7	225	2	E96742	hypothetical prote	910	6	0.7	244	2	T28175	hypothetical prote
838	6	0.7	226	1	MWGP2	25k protein - pota	911	6	0.7	245	2	S48334	GTP-binding protei
839	6	0.7	226	2	T11511	H+-transporting AT	912	6	0.7	245	2	A55456	transcription acti
840	6	0.7	226	2	C81328	hypothetical prote	913	6	0.7	246	1	A46504	chymase (EC 3.4.21
841	6	0.7	226	2	E70526	hypothetical prote	914	6	0.7	246	2	A69699	ribosomal protein
842	6	0.7	227	2	S72285	ribosomal protein	915	6	0.7	246	2	H82790	sec-independent pr
843	6	0.7	227	2	E75066	probable translati	916	6	0.7	246	2	H71548	hypothetical prote
844	6	0.7	228	2	D69627	cell-division Arp-	917	6	0.7	247	1	EVECMF	mcbf protein - Esc
845	6	0.7	228	2	AE2200	two-component resp	918	6	0.7	247	2	S59334	leucine-rich acid
846	6	0.7	229	2	S50328	cytochrome-c oxida	919	6	0.7	247	2	S67685	hypothetical prote
847	6	0.7	229	2	D38941	cytochrome-c oxida	920	6	0.7	247	2	H83970	phospho-N-acetylmu
848	6	0.7	229	2	S70599	cytochrome-c oxida	921	6	0.7	248	2	AF1399	pseudouridylylate sy
849	6	0.7	229	2	S14207	cytochrome-c oxida	922	6	0.7	248	2	C64666	glutamine transpor
850	6	0.7	229	2	E34284	cytochrome-c oxida	923	6	0.7	248	2	F71849	amino acid ABC tra
851	6	0.7	229	2	S01503	cytochrome-c oxida	924	6	0.7	248	2	F72087	conserved hypothet
852	6	0.7	229	2	T11119	cytochrome-c oxida	925	6	0.7	248	2	D86537	cr056 hypothetical
853	6	0.7	229	2	T11796	cytochrome-c oxida	926	6	0.7	248	2	T26584	hypothetical prote
854	6	0.7	229	2	T30734	hypothetical prote	927	6	0.7	248	2	F90092	hypothetical prote
855	6	0.7	230	2	E83554	probable pseudouri	928	6	0.7	248	2	C75140	hypothetical prote
856	6	0.7	230	2	C87200	conserved hypothet	929	6	0.7	248	2	H97784	hypothetical prote
857	6	0.7	230	2	S71747	DAG protein precur	930	6	0.7	249	2	JX0163	lectin II - furze
858	6	0.7	230	2	T22763	hypothetical prote	931	6	0.7	249	2	E64497	hypothetical prote
859	6	0.7	231	2	D81706	amino acid ABC tra	932	6	0.7	249	2	S43309	probable HLA class
860	6	0.7	231	2	S74286	hypothetical prote	933	6	0.7	249	2	T01891	hypothetical prote
861	6	0.7	231	2	B70795	probable cobQ2 pro	934	6	0.7	249	2	T32762	hypothetical prote
862	6	0.7	231	2	T04773	hypothetical prote	935	6	0.7	249	2	AB0537	probable membrane
863	6	0.7	231	2	S50222	hypothetical prote	936	6	0.7	249	2	A87480	ABC transporter, A
864	6	0.7	231	2	AG3225	hypothetical lysin	937	6	0.7	250	2	F83001	probable periplasm
865	6	0.7	232	2	F90088	26S proteasome SU	938	6	0.7	251	2	G69271	molycopterin oxid
866	6	0.7	232	2	T12740	hypothetical prote	939	6	0.7	251	2	A48577	filarial antigen A
867	6	0.7	232	2	T20934	hypothetical prote	940	6	0.7	252	2	S61515	dihydrodiol dehydr
868	6	0.7	232	2	T16315	hypothetical prote	941	6	0.7	252	2	S25965	H+-transporting AT
869	6	0.7	233	1	K1MST	thymidine kinase (	942	6	0.7	252	2	S25965	cyclase hisF - Rho
870	6	0.7	233	2	B69900	conserved hypothet	943	6	0.7	252	2	S54838	thyrotropin recept
871	6	0.7	233	2	S50532	hypothetical prote	944	6	0.7	252	2	T01787	probable transcrip
872	6	0.7	233	2	JC7618	FRAT2 protein - hu	945	6	0.7	252	2	S73753	probable lipoprote
873	6	0.7	234	2	S08662	dnak-type molecula	946	6	0.7	253	1	JC1319	thyrotropin recept
874	6	0.7	234	2	T15387	hypothetical prote	947	6	0.7	253	2	H72388	shikimate 5-dehydr
875	6	0.7	234	2	F86855	conserved hypothet	948	6	0.7	253	2	G87294	ATP synthase F0, A
876	6	0.7	235	2	T47426	hypothetical prote	949	6	0.7	253	2	S57586	MADS-box regulator
877	6	0.7	235	2	G97320	transcription regu	950	6	0.7	253	2	F70378	erythrocyte band 7
878	6	0.7	235	2	AD1735	hypothetical prote	951	6	0.7	253	2	G69968	hypothetical prote
879	6	0.7	236	2	S60390	probable membrane	952	6	0.7	253	2	S73473	DNA polymerase III
880	6	0.7	236	2	H95939	hypothetical prote	953	6	0.7	253	2	T31021	hypothetical prote
881	6	0.7	236	2	B83718	glycoprotein endop	954	6	0.7	254	1	C0BEQ2	cyclin homolog ECL
882	6	0.7	237	2	E71128	arginase related p	955	6	0.7	254	2	E71203	hypothetical prote
883	6	0.7	237	2	T00452	hypothetical prote	956	6	0.7	254	2	C75219	probable myo-inosi
884	6	0.7	237	2	C86309	Similar to Ubiquit	957	6	0.7	254	2	T43494	hypothetical prote
885	6	0.7	237	2	AE2037	hypothetical prote	958	6	0.7	255	2	T36085	probable short cha
886	6	0.7	238	1	TWBYA1	protein N-acetyltr	959	6	0.7	255	2	H84965	triiose-phosphate 1
887	6	0.7	238	2	S28758	H+-transporting AT	960	6	0.7	255	2	H90130	60s ribosomal prot
888	6	0.7	238	2	AH0164	arginine transport	961	6	0.7	255	2	T33401	hypothetical prote
889	6	0.7	238	2	A75607	conserved hypothet	962	6	0.7	256	2	D84037	electron transfer
890	6	0.7	238	2	T40721	hypothetical prote	963	6	0.7	256	2	S50280	PMP31 protein - ye
891	6	0.7	238	2	T04280	hypothetical prote	964	6	0.7	256	2	S50281	PMP32 protein - ye
892	6	0.7	240	2	S74395	hypothetical prote	965	6	0.7	256	2	B97565	stationary-phase s
893	6	0.7	240	2	D71822	hypothetical prote	966	6	0.7	256	2	AH2785	stationary-phase s
894	6	0.7	240	2	E84390	24-sterol C-methyl	967	6	0.7	256	2	F90403	hypothetical prote
895	6	0.7	240	2	AE1631	uroporphyrinogen I	968	6	0.7	256	2	F84612	hypothetical prote
896	6	0.7	241	1	K1BP04	hypothetical prote	969	6	0.7	256	2	T09034	hypothetical prote
897	6	0.7	241	1	D71233	deoxynucleotide mo	970	6	0.7	256	2	T46212	hypothetical prote
898	6	0.7	241	1	D71233	probable 3-isoprop	971	6	0.7	256	2	T45049	hypothetical prote
899	6	0.7	241	2	I51211	cytoskeletal prote	972	6	0.7	257	2	F86259	protein T12C24.10
900	6	0.7	241	2	A75200	hypothetical prote	973	6	0.7	257	2	G96654	hypothetical prote
901	6	0.7	242	2	G82206	hypothetical prote	974	6	0.7	257	2	F89783	hypothetical prote
902	6	0.7	243	2	J50698	hypothetical 26.6k	975	6	0.7	257	2	A82972	hypothetical prote
903	6	0.7	243	2	D81804	hypothetical prote	976	6	0.7	257	2	T40655	hypothetical prote
904	6	0.7	243	2	B81715	conserved hypothet	977	6	0.7	257	2	T18754	hypothetical prote
905	6	0.7	243	2	D71562	hypothetical prote	978	6	0.7	258	2	A45161	serine proteinase

979 6 0.7 258 2 A40360  
980 6 0.7 259 2 S36410  
981 6 0.7 259 2 A70359  
982 6 0.7 259 2 F81912  
983 6 0.7 259 2 A2167  
984 6 0.7 260 2 F97761  
985 6 0.7 260 2 B99913  
986 6 0.7 260 2 AG2760  
987 6 0.7 260 2 T24461  
988 6 0.7 260 2 D86266  
989 6 0.7 260 2 E84168  
990 6 0.7 260 2 T28182  
991 6 0.7 260 2 C98084  
992 6 0.7 261 2 B86643  
993 6 0.7 261 2 S16802  
994 6 0.7 261 2 T40773  
995 6 0.7 261 2 E59159  
996 6 0.7 261 2 C72128  
997 6 0.7 261 2 F86494  
998 6 0.7 261 2 A90661  
999 6 0.7 261 2 G64747  
1000 6 0.7 261 2 G85511

ALIGNMENTS

RESULT 1  
T05322  
hypothetical protein F18F4.240 - Arabidopsis thaliana  
N:Alternate names: hypothetical protein F1C12.60  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 31-Jan-2000  
C:Accession: T05322; T04898  
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysschaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: Z15408  
A:Accession: T05322  
A:Molecule type: DNA  
A:Residues: 1-1232 <BRV>  
A:CROSS-references: EMBL:AL022224  
A:Experimental source: cultivar Columbia; BAC clone F1C12  
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, February 1998  
A:Reference number: Z15388  
A:Accession: T04898  
A:Molecule type: DNA  
A:Residues: 1-305 <BEW>  
A:CROSS-references: EMBL:AL021637  
A:Experimental source: cultivar Columbia; BAC clone F18F4  
C:Genetics:  
A:Map position: 4  
A:Introns: 863/1; 1116/1  
A:Note: F1C12.60; F18F4.240  
C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; p

Query Match 1.5%; Score 12; DB 2; Length 1232;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 672 LETLDSHNQLT 683  
|||||  
Db 794 LETLDSHNQLT 805  
  
RESULT 2  
AC1328  
Internalin proteins homolog lmo2027 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AC1328

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
D.; Jones, L.M.; Karst, G.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurepkat, G.; Madieno, E.; Maltournam, A.;  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of *Listeria species*.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1328  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-367 <GLA>  
A:CROSS-references: GB:NC\_003210; PIDN:CAD00105.1; PID:g16411497; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo2027

Query Match 1.2%; Score 10; DB 2; Length 367;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 674 TLDLSHNQLT 683  
|||||  
Db 201 TLDLSHNQLT 210

RESULT 3  
S67265  
hypothetical protein YOR353c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein 06612  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: S67265  
R:Dellus, R.; Hebling, U.; Hofmann, B.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67261  
A:Accession: S67265  
A:Molecule type: DNA  
A:Residues: 1-791 <DEL>  
A:CROSS-references: EMBL:Z75261; NID:g1420767; PID:e252177; PID:g1420768; GSPDB:GN000  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YOR353C  
A:Map position: 15R

Query Match 1.2%; Score 10; DB 2; Length 791;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 VLSLKDNQNT 188  
|||||  
Db 141 VLSLKDNQNT 150

RESULT 4  
NBHUC8  
decorin precursor - human  
N:Alternate names: cartilage proteoglycan protein II; DS-PG II; PG40 core protein; pr  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 21-Jan-2000  
C:Accession: A45016; A45015; A26476; S05640  
R:Vetter, U.; Vogel, W.; Just, W.; Young, M.F.; Fisher, L.W.  
Genomics 15, 161-168, 1993  
A:Title: Human decorin gene: Intron-exon junctions and chromosomal localization.  
A:Reference number: A45016; MUID:93162643  
A:Accession: A45016  
A:Molecule type: DNA  
A:Residues: 1-359 <VET>  
A:CROSS-references: GB:L01125; GB:L01126; GB:L01127; GB:L01128; GB:L01129; GB:L01130;  
A:Note: sequence extracted from NCBI backbone (NCBIP:125061)  
R:Danielson, K.G.; Fazzio, A.; Cohen, I.; Cannizzaro, L.A.; Eichstetter, I.; Iozzo, R

Genomics 15, 146-160, 1993  
A:Title: The human decorin gene: intron-exon organization, discovery of two alternative  
A:Reference number: A45015; MUID:93162642  
A:Accession: A45015  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 28-70 <DA>  
A:CROSS-references: GB:M98262  
A:Note: sequence extracted from NCBI backbone (NCBIP:125013)  
A:Accession: B45015  
A:Molecule type: DNA  
A:Residues: 296-359 <DAN>  
A:Note: sequence extracted from NCBI backbone (NCBIP:125017)  
R:Krusius, T.; Ruoslahti, E.  
Proc. Natl. Acad. Sci. U.S.A. 83, 7683-7687, 1986  
A:Title: Primary structure of an extracellular matrix proteoglycan core protein deduced  
A:Reference number: A26476; MUID:87017013  
A:Accession: A26476  
A:Molecule type: mRNA  
A:Residues: 1-359 <KRU>  
A:CROSS-references: GB:M14219; NID:g181169; PIDN:AAB00774.1; PID:g181170  
R:Roughley, P.J.; White, R.J.  
Biochem. J. 262, 823-827, 1989  
A:Title: Dermatan sulphate proteoglycans of human articular cartilage. The properties of  
A:Reference number: S05639; MUID:90073579  
A:Accession: S05640  
A:Molecule type: protein  
A:Residues: 31-33,'X',35-50 <ROU>  
C:Comment: This protein binds type I collagen.  
C:Genetics:  
A:Gene: GDB:DCN  
A:CROSS-references: GDB:119839; OMIM:125255  
A:Map position: 12q21.3-12q23  
A:Introns: 71/1; 108/3; 180/1; 218/1; 249/2; 295/3  
A:Note: the first two introns occur before the initiator codon  
C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan  
C:Keywords: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; dupli  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-30/Domain: propeptide #status predicted <PRO>  
F:31-359/Product: decorin #status predicted <MPT>  
F:48-72/Domain: proteoglycan amino-terminal homology <PAH>  
F:82-105/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:106-129/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:130-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:151-174/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:175-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:201-221/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:222-245/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:246-269/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:270-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:308-359/Domain: proteoglycan carboxyl-terminal homology <PRC>  
F:34/Binding site: dermatan sulfate (Ser) (covalent) #status experimental  
F:189,325/Binding site: dermatan sulfate (Ser) (covalent) #status predicted  
F:211,262,303/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.1%; Score 9; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121  
|||||||

DB 215 IPQGLPPSL 223

RESULT 5  
S06280

decorin precursor - bovine

N:Alternate names: dermatan sulfate proteoglycan II; proteoglycan core protein II

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 20-Aug-1999

C:Accession: S06280; B31430; A26545; A20935  
R:Day, A.A.; McQuillan, C.I.; Terimine, J.D.; Young, M.R.  
Biochem. J. 248, 801-805, 1987  
A:Title: Molecular cloning and sequence analysis of the cDNA for small proteoglycan I  
A:Reference number: S06280; MUID:88133946  
A:Accession: S06280  
A:Molecule type: mRNA  
A:Residues: 1-360 <DAY>  
A:CROSS-references: EMBL:Y00712; NID:g618; PIDN:CAA68702.1; PID:g619  
A:Experimental source: bone  
R:Choi, H.U.; Johnson, T.L.; Pal, S.; Tang, L.H.; Rosenberg, L.; Neame, P.J.  
J. Biol. Chem. 264, 2876-2884, 1989  
A:Title: Characterization of the dermatan sulfate proteoglycans of bovine sclera and their relationship t  
A:Reference number: A26545; MUID:87137687  
A:Accession: A26545  
A:Molecule type: protein  
A:Residues: 31-33,'X',35-54 <CHO>  
A:Experimental source: cartilage; fetal skin  
R:Coster, L.; Rosenberg, L.C.; van der Rest, M.; Poole, A.R.  
J. Biol. Chem. 262, 3809-3812, 1987  
A:Title: The dermatan sulfate proteoglycans of bovine sclera and their relationship t  
A:Reference number: A26545; MUID:87137687  
A:Accession: A26545  
A:Molecule type: protein  
A:Residues: 31-50 <COS>  
A:Experimental source: sclera  
R:Pearson, C.H.; Winterbottom, N.; Fackre, D.S.; Scott, P.G.; Carpenter, M.R.  
J. Biol. Chem. 258, 15101-15104, 1983  
A:Reference number: A20935; MUID:84087911  
A:Accession: A20935  
A:Molecule type: protein  
A:Residues: 31-54 <PEA>  
A:Experimental source: skin  
R:Chopra, R.K.; Pearson, C.H.; Pringle, G.A.; Fackre, D.S.; Scott, P.G.  
Biochem. J. 232, 277-279, 1985  
A:Title: Dermatan sulphate is located on serine-4 of bovine skin proteodermatan sulph  
nces around glycosylation sites in different proteoglycans.  
A:Reference number: A44700; MUID:86103195  
A:Contents: annotation; glycosylation  
C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglyc  
C:Keywords: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; ext  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-30/Domain: propeptide #status predicted <PRO>  
F:31-360/Product: decorin #status predicted <MAT>  
F:49-73/Domain: proteoglycan amino-terminal homology <PAH>  
F:83-106/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:107-130/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:131-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:152-175/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:176-199/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:202-222/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:223-246/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:247-270/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:271-293/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:294-308/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:309-360/Domain: proteoglycan carboxyl-terminal homology <PCH>  
F:34/Binding site: dermatan sulfate (Ser) (covalent) #status experimental  
F:190,326/Binding site: dermatan sulfate (Ser) (covalent) #status predicted  
F:212,263,304/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.1%; Score 9; DB 2; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121  
|||||||

DB 216 IPQGLPPSL 224

RESULT 6  
147020

decorin - rabbit



A:Cross-references: FlyBase:FBgn0000313  
A:Introns: 1/3 80/3; 318/3; 377/2 422/2; 702/1 745/3; 831/2; 998/2  
C:Superfamily: chaoptin; leucine-rich alpha-2-glycoprotein repeat homology  
C:Keywords: cell adhesion; glycoprotein; membrane protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-1134/Product: chaoptin #status predicted <MAT>  
F:180-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR1>  
F:103-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>  
F:128-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>  
F:152-175/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>  
F:177-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>  
F:201-224/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>  
F:226-249/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>  
F:250-273/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>  
F:279-302/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>  
F:303-325/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>  
F:326-349/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>  
F:351-374/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>  
F:375-399/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
F:401-424/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
F:428-451/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
F:453-476/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
F:477-500/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>  
F:502-526/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>  
F:527-550/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>  
F:551-574/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR20>  
F:577-600/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR21>  
F:601-624/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR22>  
F:625-648/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR23>  
F:649-672/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR24>  
F:673-696/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR25>  
F:708-731/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR26>  
F:733-756/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR27>  
F:757-780/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR28>  
F:781-804/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR29>  
F:805-827/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR30>  
F:828-851/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR31>  
F:854-877/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR32>  
F:879-902/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR33>  
F:903-926/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR34>  
F:928-948/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR35>  
F:949-972/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR36>  
F:973-995/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR37>  
F:996-1019/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR38>  
F:1021-1044/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR39>  
F:1056-1080/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR40>

Query Match 1.1%; Score 9; DB 1; Length 1134;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 222 LQILDLSGN 230  
DB 352 LQILDLSGN 360  
IIIIIIIIII  
RESULT 9  
A49674  
flightless-I homolog - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: A49674  
R:Campbell, H.D.; Schlumsky, T.; Claudianos, C.; Ozsarac, N.; Kasprzak, A.B.; Cotseil,  
Proc. Natl. Acad. Sci. U.S.A. 90, 11386-11390, 1993  
A:Title: The Drosophila melanogaster flightless-I gene involved in gastrulation and musc  
s and humans.  
A:Reference number: A49674; MUID:94068608  
A:Accession: A49674  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1268 <RES>  
A:Cross-references: EMBL:U01184; NID:g440176; PIDN:AAC03568.1; PID:g440177

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; gelsolin repeat hom  
F:498-838/Domain: gelsolin repeat homology <GEL1>  
F:904-1361/Domain: gelsolin repeat homology <GEL2>  
Query Match 1.1%; Score 9; DB 2; Length 1268;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 675 LDLSHNQLT 683  
DB 108 LDLSHNQLT 116  
IIIIIIIIII  
RESULT 10  
S56221  
hypothetical protein YFL033c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 02-Sep-1995 #sequence\_revision 19-Oct-1995 #text\_change 29-Oct-1999  
C:Accession: S56221  
R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas  
submitted to the EMBL Data Library, May 1995  
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces  
A:Reference number: S56186  
A:Accession: S56221  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1770 <MUR>  
A:Cross-references: EMBL:D50617; NID:g836685; PIDN:BAA09206.1; PID:d1009846; PID:g836  
C:Genetics:  
A:Gene: SGD:RIM15  
A:Cross-references: SGD:S0001861; MIPS:YFL033c  
A:Map position: 6L

Query Match 1.1%; Score 9; DB 2; Length 1770;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 646 PNLKNLSLA 654  
DB 1397 PNLKNLSLA 1405  
IIIIIIIIII  
RESULT 11  
C96615  
hypothetical protein T18124.10 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96615  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: C96615  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1784 <SPO>  
A:Cross-references: GB:AE005173; NID:g11038494; PIDN:AAG27771.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T18124.10  
A:Map position: 1  
Query Match 1.1%; Score 9; DB 2; Length 1784;  
Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 LOELDLSON 295

|||||

Db 226 LOELDLSON 234

#### RESULT 12

S10005

hypothetical protein 5 - fowl adenovirus 1

C:Species: Avidenovirus galli (fowl adenovirus 1, CELO)

C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 20-Apr-2000

C:Accession: S10005

R:Akopian, T.A.; Kruglyak, V.A.; Rivkina, M.B.; Naroditsky, B.S.; Tikhonenko, T.I.

Nucleic Acids Res. 18, 2825, 1990

A:Title: Sequence of an avian adenovirus (CELO) DNA fragment (0-11.2%).

A:Reference number: S10004; MUID:90251474

A:Accession: S10005

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-283 <AKO>

A:Cross-references: EMBL:X17217; NID:g58537; PIDN:CAA35087.1; PID:g58539

C:Superfamily: fowl adenovirus 1 hypothetical protein 5

#### Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 283;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 TNLTLTN 46

|||||

Db 222 TNLTLTN 229

#### RESULT 13

A10810

pyridoxal kinase (EC 2.7.1.35) - Salmonella enterica subsp. enterica serovar Typhi (stra

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001

C:Accession: A10810

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Farrar, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: A10810

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-288 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD07667.1; PID:gl6503653; GSPDB:GN00176

C:Genetics:

A:Gene: STY2672

C:Keywords: phosphotransferase

#### Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 288;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 VTAVPTVL 194

|||||

Db 51 VTAVPTVL 58

#### RESULT 14

T49908

hypothetical protein T24H18.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49908

R:Bevan, M.; Robben, J.; Grymonprez, B.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Ru  
submitted to the Protein Sequence Database, April 2000

A:Reference number: 225024

A:Accession: T49908

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-371 <BEV>

A:Cross-references: EMBL:AL353013; GSPDB:GN00063; ATSP:T24H18.110

A:Experimental source: cultivar Columbia; BAC clone T24H18

C:Genetics:

A:Gene: ATSP:T24H18.110

A:Map position: 5

#### Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 371;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 493 LNLGNLI 500

|||||

Db 282 LNLGNLI 289

#### RESULT 15

T51702

amidophosphoribosyltransferase (EC 2.4.2.14) [similarity] - Lactococcus lactis

N:Alternate names: phosphoribosylpyrophosphate amidotransferase

C:Species: Lactococcus lactis

C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 03-Aug-2001

C:Accession: T51702

R:Peltonen, T.; Mantasala, P.

Mol. Gen. Genet. 261, 31-41, 1999

A:Title: Isolation and characterization of a purC(orf)QLF operon from Lactococcus lac

A:Reference number: 225429; MUID:99168765

A:Accession: T51702

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-506 <PEL>

A:Cross-references: EMBL:U64311; PIDN:MDI2627.1

A:Experimental source: strain MGI614

C:Genetics:

A:Gene: purF

C:Superfamily: amidophosphoribosyltransferase

C:Keywords: glycosyltransferase; pentosyltransferase; purine nucleotide biosynthesis

F:46/Active site: Cys #status predicted

#### Query Match

Best Local Similarity 100.0%; Score 8; DB 2; Length 506;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 SLAKNGLK 659

|||||

Db 499 SLAKNGLK 506

#### RESULT 16

S26857

isocitrate lyase (EC 4.1.3.1) - Emericella nidulans

N:Alternate names: isocitrase; isocitratase; isocitritase

C:Species: Emericella nidulans, Aspergillus nidulans

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 23-Mar-2001

C:Accession: S26857; S22055

R:Ganey, L.D.S.; Connerton, I.F.; Lewis, E.H.; Turner, G.; Ballance, D.J.

Curr. Genet. 21, 43-47, 1992

A:Title: Characterization of the glyoxysomal isocitrate lyase genes of Aspergillus ni

A:Reference number: S26857; MUID:92136435

A:Accession: S26857

A:Molecule type: DNA

A:Residues: 1-537 <GAI>

A:Cross-references: EMBL:X62696

R:Connerton, I.F.

submitted to the EMBL Data Library, October 1991

A;Description: Characterization of the glyoxysomal isocitrate lyase genes of *Aspergillus*  
A;Reference number: S22055  
A;Accession: S22055  
A;Molecule type: DNA  
A;Residues: 1-66, 68-537 <CON>  
A;Cross-references: EMBL:X62696; NID:g2316; PIDN:CAA44572.1; PID:g2317  
C;Genetics:  
A;Gene: acud  
A;Introns: 67/3; 119/3  
A;Superfamily: isocitrate lyase  
C;Keywords: carbon-carbon lyase; glyoxylate bypass; glyoxysome; oxo-acid-lyase

Query Match 1.0%; Score 8; DB 2; Length 537;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 448 FKNKEASF 455  
Db 66 FKNKEASF 73  
|||||

RESULT 17  
E96598  
protein F20N2.2 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: E96598  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719  
A;Accession: E96598  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-607 <STO>  
A;Cross-references: GB:AE005173; NID:g8778504; PIDN:AAF79512.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: F20N2.2  
A;Map position: 1

Query Match 1.0%; Score 8; DB 2; Length 607;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 NLEVLIDG 376  
Db 389 NLEVLIDG 396  
|||||

RESULT 18  
T10727  
protein kinase Xa21 (EC 2.7.1.-) D, receptor type - long-staminate rice  
C;Species: Oryza longistaminata (long-staminate rice)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C;Accession: T10727  
R;Song, W.Y.; Pi, L.Y.; Wang, G.L.; Gardner, J.; Holsten, T.; Ronald, P.C.  
Plant Cell 9, 1279-1287, 1997  
A;Title: Evolution of the rice Xa21 disease resistance gene family.  
A;Reference number: Z15276; MUID:97432142  
A;Accession: T10727  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-612 <SON>  
A;Cross-references: EMBL:U72726; NID:g2586078; PIDN:AAB82753.1; PID:g2586081

A;Description: Characterization of the glyoxysomal isocitrate lyase genes of *Aspergillus*  
A;Reference number: S22055  
A;Accession: S22055  
A;Molecule type: DNA  
A;Residues: 1-66, 68-537 <CON>  
A;Cross-references: EMBL:X62696; NID:g2316; PIDN:CAA44572.1; PID:g2317  
C;Genetics:  
A;Gene: acud  
A;Introns: 67/3; 119/3  
A;Superfamily: isocitrate lyase  
C;Keywords: carbon-carbon lyase; glyoxylate bypass; glyoxysome; oxo-acid-lyase

Query Match 1.0%; Score 8; DB 2; Length 537;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 448 FKNKEASF 455  
Db 66 FKNKEASF 73  
|||||

RESULT 17  
E96598  
protein F20N2.2 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: E96598  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719  
A;Accession: E96598  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-607 <STO>  
A;Cross-references: GB:AE005173; NID:g8778504; PIDN:AAF79512.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: F20N2.2  
A;Map position: 1

Query Match 1.0%; Score 8; DB 2; Length 607;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 NLEVLIDG 376  
Db 389 NLEVLIDG 396  
|||||

RESULT 18  
T10727  
protein kinase Xa21 (EC 2.7.1.-) D, receptor type - long-staminate rice  
C;Species: Oryza longistaminata (long-staminate rice)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C;Accession: T10727  
R;Song, W.Y.; Pi, L.Y.; Wang, G.L.; Gardner, J.; Holsten, T.; Ronald, P.C.  
Plant Cell 9, 1279-1287, 1997  
A;Title: Evolution of the rice Xa21 disease resistance gene family.  
A;Reference number: Z15276; MUID:97432142  
A;Accession: T10727  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-612 <SON>  
A;Cross-references: EMBL:U72726; NID:g2586078; PIDN:AAB82753.1; PID:g2586081

A;Experimental source: strain IRBB21  
C;Genetics:  
A;Map position: 11  
C;Keywords: phosphotransferase

Query Match 1.0%; Score 8; DB 2; Length 612;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682  
Db 156 LDLSHNQL 163  
|||||

RESULT 19  
S47299  
gene F protein - rinderpest virus  
C;Species: rinderpest virus  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
C;Accession: S47299  
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A;Description: The complete nucleotide sequence of the fusion protein gene of the vac  
A;Reference number: S47299  
A;Accession: S47299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-636 <EVA>  
A;Cross-references: EMBL:Z31655; NID:g535391; PIDN:CAA83481.1; PID:g535392  
C;Superfamily: parainfluenza virus cell fusion protein

Query Match 1.0%; Score 8; DB 2; Length 636;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 774 VTIPYLAT 781  
Db 101 VTIPYLAT 108  
|||||

RESULT 20  
C84633  
probable disease resistance protein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: C84633  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487  
A;Accession: C84633  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-743 <STO>  
A;Cross-references: GB:AE002093; NID:g4115376; PIDN:AAD03377.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g24160  
A;Map position: 2

Query Match 1.0%; Score 8; DB 2; Length 743;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682  
Db 608 LDLSHNQL 615  
|||||





A;Residues: 1-853 <PAR>  
A;Cross-references: EMBL:AF119040; NID:g4235640; PID:g4235641; PIDN:AAD13301.1  
C;Genetics:  
A;Gene: NL0D

Query Match 1.0%; Score 8; DB 2; Length 853;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSSNKI 730  
|||||  
DB 714 LDLSSNKI 721

## RESULT 26

T07015  
Cf-4A protein - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 11-May-2000  
C;Accession: T07015  
R;Takken, F.L.; Schipper, D.; Nijkamp, H.J.; Hill, J.  
Plant J. 14, 401-411, 1998  
A;Title: Identification and Ds-tagged isolation of a new gene at the Cf-4 locus of tomato  
A;Reference number: Z15863; MUID:98335213  
A;Accession: T07015  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-855 <TAK>  
A;Cross-references: EMBL:Y12640; NID:el289424; PIDN:CAA73187.1; PID:el289425  
A;Experimental source: strain Cf-4; isolate MM-Cf-4  
C;Genetics:  
A;Gene: Cf-4A  
A;Map position: 1  
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

Query Match 1.0%; Score 8; DB 2; Length 855;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSSNKI 730  
|||||  
DB 713 LDLSSNKI 720

## RESULT 27

T17460  
disease resistance protein - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000  
C;Accession: T17460  
R;Parniske, M.; Jones, J.D.  
Proc. Natl. Acad. Sci. U.S.A. 96, 5850-5855, 1999  
A;Title: Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene  
A;Reference number: Z18801; MUID:99254130  
A;Accession: T17460  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-855 <PAR>  
A;Cross-references: EMBL:AF119040; NID:g4235640; PID:g4235642; PIDN:AAD13302.1  
C;Genetics:  
A;Gene: NL0C

Query Match 1.0%; Score 8; DB 2; Length 855;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSSNKI 730  
|||||  
DB 713 LDLSSNKI 720

## RESULT 28

A55173  
Cf-9 protein precursor - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 20-Jan-1995 #sequence\_revision 20-Jan-1995 #text\_change 21-Jan-2000  
C;Accession: A55173  
R;Jones, D.A.; Thomas, C.M.; Hammond-Kosack, K.E.; Balint-Kurti, P.J.; Jones, J.D.G.  
Science 266, 789-793, 1994  
A;Title: Isolation of the tomato Cf-9 gene for resistance to Cladosporium fulvum by transposon activation  
A;Reference number: A55173; MUID:95063912  
A;Accession: A55173  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-863 <JON>  
A;Cross-references: GB:U15936; NID:g563232; PID:g563233  
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology  
F;717-739/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 1.0%; Score 8; DB 2; Length 863;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSSNKI 730  
|||||  
DB 721 LDLSSNKI 728

## RESULT 29

B96770  
hypothetical protein F1017.14 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: B96770  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, U.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, D.; Narsisu, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719  
A;Accession: B96770  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-910 <STO>  
A;Cross-references: GB:AE005173; NID:g6939235; PIDN:AAF31737.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: F1017.14  
A;Map position: 1

Query Match 1.0%; Score 8; DB 2; Length 910;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 QELDSLQN 295  
|||||  
DB 192 QELDSLQN 199

## RESULT 30

T50850  
receptor protein kinase homolog [imported] - soybean  
C;Species: Glycine max (soybean)  
C;Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000  
C;Accession: T50850  
R;Yamamoto, E.; Karakaya, H.C.; Knap, H.T.  
Biochim. Biophys. Acta 1491, 333-340, 2000

A:Title: Molecular characterization of two soybean homologs of Arabidopsis thaliana CLAV  
A:Reference number: 225262  
A:Accession: T50850  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-987 <YAM>  
A:Cross-references: EMBL:AF197947; PIDN:AAF59906.1  
C:Genetics:  
A:Gene: CLVIB  
C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; F

Query Match 1.0%; Score 8; DB 2; Length 987;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 720 LRYLDLS 727  
|||||||  
DB 245 LRYLDLS 252

RESULT 31  
A57676  
Protein kinase Xa21 (EC 2.7.1.-), receptor type precursor - rice  
C:Species: Oryza sativa (rice)  
C:Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 11-Jun-1999  
C:Accession: A57676  
R:Song, W.Y.; Wang, G.L.; Chen, L.L.; Kim, H.S.; Pi, L.Y.; Holsten, T.; Gardner, J.; Wan  
Science 270, 1804-1806, 1995  
A:Title: A receptor kinase-like protein encoded by the rice disease resistance gene, Xa2  
A:Reference number: A57676; MUID:96106403  
A:Accession: A57676  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-1025 <SON>  
A:Cross-references: GB:U37133; NID:g1122442; PIDN:AAC49123.1; PID:g1122443  
C:Genetics:  
A:Gene: Xa21  
A:Note: confers resistance against the bacterial pathogen Xanthomonas oryzae pv. oryzae  
C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; F  
C:Keywords: ATP; glycoprotein; magnesium; phosphotransferase; receptor; serine/threonine  
F:1-23/Domain: signal sequence #status predicted <Sig>  
F:24-1025/Product: protein kinase Xa21 #status predicted <MAT>  
F:80-103/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:104-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:128-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:153-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:177-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:201-224/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:225-248/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:249-273/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:274-297/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:298-321/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:322-345/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
F:352-375/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
F:377-400/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>  
F:401-424/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
F:425-448/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>  
F:449-472/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>  
F:473-496/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>  
F:498-521/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>  
F:522-545/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>  
F:546-569/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR20>  
F:570-593/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR21>  
F:594-618/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR22>  
F:706-1011/Domain: protein kinase homology <Kin>  
F:714-722/Region: protein kinase ATP-binding motif  
F:55-90-101,198,235,246,295,322,349,373,435,446,470,483,503,580,599/Binding site: carbo  
F:736,752,842,844/Active site: Lys, Glu, Asp, Lys #status predicted  
F:847,851/Binding site: magnesium (Asp, Asp) #status predicted

Query Match

1.0%; Score 8; DB 1; Length 1025;

Best Local Similarity 100.0%; Pred. No. 37;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682  
|||||||  
DB 156 LDLSHNQL 163

RESULT 32  
H70203  
Isoleucine--trNA ligase (EC 6.1.1.5) lles - Lyme disease spirochete  
N:Alternate names: isoleucyl-trNA synthetase  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 16-Jul-1999  
C:Accession: H70203  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: H70203  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1042 <KLE>  
A:Cross-references: GB:AE001181; GB:AE000783; NID:g2688772; PIDN:AAC67179.1; PID:g268  
A:Experimental source: Strain B31  
C:Superfamily: isoleucine--trNA synthetase  
C:Keywords: aminoacyl-trNA synthetase; ligase; protein biosynthesis

Query Match 1.0%; Score 8; DB 2; Length 1042;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 KNLSLAKN 656  
|||||||  
DB 670 KNLSLAKN 677

RESULT 33

B86479

hypothetical protein F14D7.1 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C:Accession: B86479

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marzla  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: B86479

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1120 &lt;STO&gt;

A:Cross-references: GB:AE005172; NID:g8778966; PIDN:AAF79881.1; GSPDB:GN00141  
C:Genetics:

A:Map position: 1

Query Match 1.0%; Score 8; DB 2; Length 1120;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682  
|||||||

Db 682 LDLSHNQL 689

## RESULT 34

T10636

hypothetical protein T13K14.100 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999

C:Accession: T10636

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16991

A:Accession: T10636

A:Molecule type: DNA

A:Residues: 1-1143 <BEV>

A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.100

A:Experimental source: cultivar Columbia; BAC clone T13K14

C:Genetics:

A:Gene: ATSP:T13K14.100

A:Map position: 4

A:Introns: 334/1; 815/2; 941/3; 970/2; 1006/3; 1031/1; 1073/3; 1105/3

## Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 1143;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 VLDISSNS 548

|||||

Db 439 VLDISSNS 446

## RESULT 35

AE1852

hypothetical protein alr0366 [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002

C:Accession: AE1852

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE1852

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1152 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA072324.1; PID:g17129711; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr0366

## Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 1152;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 IEKDAFLN 173

|||||

Db 804 IEKDAFLN 811

## RESULT 36

C84568

hypothetical protein Atg18760 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: C84568

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

RESULT 38

B86398

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487

A:Accession: C84568

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1187 <STO>

A:Cross-references: GB:AE002093; NID:g4185142; PIDN:AA08945.1; GSPDB:GN00139

C:Genetics:

A:Gene: Atg18760

A:Map position: 2

## Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 1187;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 QLFKNLLK 622

|||||

Db 1158 QLFKNLLK 1165

## RESULT 37

S54587

CAT8 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: MSP8 protein; protein YMR021.06c; protein YMR280c

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 24-Sep-1999

C:Accession: S54587; S48234; S61595; S49498

R:Pearson, D.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54587

A:Accession: S54587

A:Molecule type: DNA

A:Residues: 1-1433 <PEA>

A:Cross-references: EMBL:Z49704; NID:g825540; PIDN:CAA89778.1; PID:g825546; MIPS:YMR2

A:Experimental source: strain AB972

R:Grzesitz, D.

submitted to the EMBL Data Library, March 1994

A:Reference number: S48234

A:Accession: S48234

A:Molecule type: DNA

A:Residues: 1-746; 'L', 748-1433 <GR2>

A:Cross-references: EMBL:X78344; NID:g559523; PIDN:CAA55139.1; PID:g559524

R:Boles, E.; Hettmann, C.; Zimmermann, F.K.

submitted to the EMBL Data Library, December 1995

A:Reference number: S61594

A:Accession: S61595

A:Molecule type: DNA

A:Residues: 1-302; 'A', 304-767; 'A', 769-998; 1003-1007; 'S', 1009-1015; 'Q', 1017-1018; 'P', 1

A:Cross-references: EMBL:X94215; NID:g1122900; PIDN:CAA63906.1; PID:e214033; PID:g112

A:Experimental source: strain ENY.WA-1A

C:Genetics:

A:Gene: SGD:CAT8; MSP8

A:Cross-references: SGD:S0004893; MIPS:YMR280c

A:Map position: 13R

C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluste

C:Keywords: transmembrane protein

F:65-102/Domain: GAL4 zinc binuclear cluster homology <GAL4>

F:456-472/Domain: transmembrane #status predicted <TM1>

F:738-754/Domain: transmembrane #status predicted <TM2>

## Query Match

Best Local Similarity 1.0%; Score 8; DB 2; Length 1433;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 PSLQLLSL 126

|||||

Db 520 PSLQLLSL 527

protein T7N9.24 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: B86398  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: B86398  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1590 <STO>  
 A:Cross-references: GB:AE005172; NID:g10121909; PIDN:AAG13419.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: T7N9.24  
 A:Map position: 1

Query Match 1.0%; Score 8; DB 2; Length 1590;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 620 LKLEELD 627  
 DB 1269 LKLEELD 1276  
 |||||

RESULT 39  
 F71704

ribosomal protein L36 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 17-Nov-2000  
 C:Accession: F71704  
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark,  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71630; MUID:99039499  
 A:Accession: F71704  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-41 <AND>  
 A:Cross-references: GB:A7235271; GB:A7235269; NID:g3868717; PIDN:CAA14912.1; PID:g386101  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: rpmJ; RP456  
 C:Superfamily: Escherichia coli ribosomal protein L36

Query Match 0.9%; Score 7; DB 2; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 SSKSLK 344  
 DB 5 SSKSLK 11  
 |||||

RESULT 40  
 C87660

ribosomal protein L36 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: G87660  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: G87660  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-41 <STO>  
 A:Cross-references: GB:AE005673; NID:g13425019; PIDN:AAK25283.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC3321

Query Match 0.9%; Score 7; DB 2; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 SSKSLK 344  
 DB 5 SSKSLK 11  
 |||||

RESULT 41  
 S28769

hypothetical protein 41 (psbI 3' region) - barley chloroplast  
 C:Species: chloroplast Hordeum vulgare (barley)  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 21-Jul-2000  
 C:Accession: S28769  
 R:Sexton, T.B.; Jones, J.T.; Mullet, J.E.  
 Curr. Genet. 17, 445-454, 1990  
 A:Title: Sequence and transcriptional analysis of the barley ctDNA region upstream of  
 A:Reference number: S28765; MUID:90291518  
 A:Accession: S28769  
 A:Molecule type: DNA  
 A:Residues: 1-41 <SEX>  
 A:Cross-references: EMBL:X52765; NID:g11601; PIDN:CAA36976.1; PID:g11606  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast

Query Match 0.9%; Score 7; DB 2; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 SKNSIFF 478  
 DB 11 SKNSIFF 17  
 |||||

RESULT 42  
 B97785

50S ribosomal protein L36 [imported] - Rickettsia conorii (strain Malish 7)  
 C:Species: Rickettsia conorii  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: B97785  
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;  
 Science 293, 2093-2098, 2001  
 A:Title: Mechanisms of evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: B97785  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-41 <KUR>  
 A:Cross-references: GB:AE006914; PIDN:AAL03220.1; PID:gl5619772; GSPDB:GN00173  
 C:Genetics:  
 A:Gene: rpmJ

Query Match 0.9%; Score 7; DB 2; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 SSKSLK 344

Db 5 SSLSLK 11  
|||||||

## RESULT 43

S35170

Cytochrome P450 (clone 5) - Madagascar periwinkle (fragment)  
N:Contains: oxidoreductase (EC 1.-.-.-)  
C:Species: Catharanthus roseus (Madagascar periwinkle)  
C:Date: 13-Jan-1995 #sequence\_revision 09-May-1997 #text\_change 04-Mar-2000  
C:Accession: S35170  
R:Meijer, A.H.; Souer, B.; Verpoorte, R.; Hoge, J.H.C.  
Plant Mol. Biol. 22, 379-383, 1993

A:Title: Isolation of cytochrome P-450 cDNA clones from the higher plant Catharanthus roseus  
A:Reference number: S35168; MUID:93283641

A:Accession: S35170

A:Molecule type: mRNA

A:Residues: 1-65 &lt;MEI&gt;

A:Cross-references: EMBL:X69777; NID:g395305; PIDN:CAA49432.1; PID:g395306

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;  
F:2/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.9%; Score 7; DB 2; Length 65;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 621 LKLELD 627

|||||||

Db 31 LKLELD 37

## RESULT 44

AG0276

Phage hypothetical protein YPO2271 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AG0276

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0276

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-87 &lt;KUR&gt;

A:Cross-references: GB:AL590842; PIDN:CAC91075.1; PID:gl5980266; GSPDB:GN00175

C:Genetics:

A:Gene: YPO2271

Query Match

Best Local Similarity 0.9%; Score 7; DB 2; Length 87;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LTKLKVL 180

|||||||

Db 81 LTKLKVL 87

## RESULT 45

B71391

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - common lancelet mitochondrion

C:Species: mitochondrion Branchiostoma lanceolatum (common lancelet)

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000

C:Accession: B71391

R:Spruyt, N.; Delarbre, C.; Gachelin, G.; Laudet, V.

Nucleic Acids Res. 26, 3279-3285, 1998

A:Title: Complete sequence of the amphioxus (Branchiostoma lanceolatum) mitochondrial ge

A:Reference number: A71390; MUID:98292550

A:Accession: B71391

A&gt;Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-91 &lt;SPR&gt;

A:Cross-references: GB:Y16474; NID:g3292989; PIDN:CAA76255.1; PID:g3292999

C:Genetics:

A:Gene: NADH4L

A:Genome: mitochondrion

A:Genetic code: SGC4

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation

Query Match

Best Local Similarity 0.9%; Score 7; DB 2; Length 91;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 LKSPNLS 362

|||||||

Db 85 LKSPNLS 91

## RESULT 46

AD0406

arsenical resistance operon repressor [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001

C:Accession: AD0406

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0406

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-113 &lt;KUR&gt;

A:Cross-references: GB:AL590842; PIDN:CAC92576.1; PID:gl5981273; GSPDB:GN00175

C:Genetics:

A:Gene: arsR

C:Superfamily: arsenical resistance operon repressor

Query Match

Best Local Similarity 0.9%; Score 7; DB 2; Length 113;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 614 LQLFKNL 620

|||||||

Db 7 LQLFKNL 13

## RESULT 47

A23329

Ly-5-8 glycoprotein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 23-Jul-1999

C:Accession: A23329

R:Shen, F.W.; Saga, Y.; Litman, G.; Freeman, G.; Tung, J.S.; Cantor, H.; Boyse, E.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 7360-7363, 1985

A:Reference number: A23329; MUID:86042665

A:Accession: A23329

A:Molecule type: mRNA

A:Residues: 1-115 &lt;SHE&gt;

A:Cross-references: GB:M11934; NID:gl98919; PIDN:AAA39461.1; PID:gl98920

C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain ho

C:Keywords: glycoprotein

Query Match

Best Local Similarity 0.9%; Score 7; DB 2; Length 115;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 740 ENVLNNL 746
      |||||
Db 91 ENVLNNL 97

RESULT 48
BVECAR
arsenical resistance operon repressor - Escherichia coli plasmid R773
C:Species: Escherichia coli
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: JS0448
R:San Francisco, M.J.D.; Hope, C.L.; Owolabi, J.B.; Tisa, L.S.; Rosen, B.P.
Nucleic Acids Res. 18, 619-624, 1990
A:Title: Identification of the metalloregulatory element of the plasmid-encoded arsenic
A:Reference number: JS0448; MUID:90174986
A:Accession: JS0448
A:Molecule type: DNA
A:Residues: 1-117 <SAN>
A:Cross-references: GB:X16045; NID:g42716; PIDN:CAA34168.1; PID:g42717
C:Comment: This is a transcriptional repressor for the ars operon; it is a trans-acting
C:Genetics:
A:Gene: arsR
A:Genome: plasmid
C:Superfamily: arsenical resistance operon repressor
C:Keywords: DNA binding; homodimer; repressor; transcription regulation

Query Match 0.9%; Score 7; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 614 LQLFKNL 620
      |||||
Db 7 LQLFKNL 13

RESULT 49
S64446
probable membrane protein YGR137w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G6413
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C:Accession: S64446
R:Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Tallia, E.; Nawrocki, A.; Del
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64428
A:Accession: S64446
A:Molecule type: DNA
A:Residues: 1-124 <VAN>
A:Cross-references: EMBL:Z72921; NID:g1323226; PID:g1323228; GSPDB:GN00007; MIPS:YGR137w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YGR137w
A:Map position: 7R
C:Superfamily: Saccharomyces probable membrane protein YGR137w
C:Keywords: transmembrane protein
F:75-91/Domain: transmembrane #status predicted <TM>

Query Match 0.9%; Score 7; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 LRLHSNS 272
      |||||
Db 10 LRLHSNS 16

RESULT 50
S31006
gene 61 protein - Mycobacterium phage L5
C:Species: Mycobacterium phage L5
```

```
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: S31006
R:Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
Mol. Microbiol. 7, 407-417, 1993
A:Title: Superinfection immunity of mycobacteriophage L5: applications for genetic tr
A:Reference number: S30949; MUID:93211283
A:Accession: S31006
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-125 <DON>
A:Cross-references: EMBL:Z18946; NID:g15859; PIDN:CAA79437.1; PID:g15917
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
C:Genetics:
A:Gene: 61

Query Match 0.9%; Score 7; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 RSFSGLT 98
      |||||
Db 28 RSFSGLT 34

Search completed: July 17, 2002, 09:45:33
Job time: 85 sec
```

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2002, 09:45:38 ; Search time 38.03 Seconds  
(without alignments)  
3670.968 Million cell updates/sec

Title: US-09-202-054-l\_COPY\_30\_836

Perfect score: 807

Sequence: 1 FPKTLPDVTLDVPMNHVIV.....PGAHRGQSVISLDLYTCELD 807

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 6

Total number of hits satisfying chosen parameters: 6621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- SPTREMBL19:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	807	100.0	1049	4 Q9NYK1	Q9nyk1 homo sapien
2	695	86.1	1049	4 Q9NR98	Q9nr98 homo sapien
3	34	4.2	1050	11 Q923I1	Q923i1 mus musculus
4	17	2.1	1032	11 Q91X17	Q91x17 mus musculus
5	12	1.5	363	4 Q9H5G9	Q9h5g9 homo sapien
6	12	1.5	1232	10 Q9SN91	Q9sn91 arabidopsis
7	10	1.2	791	3 Q988I7	Q988i7 saccharomyc
8	10	1.2	1041	4 Q9NR97	Q9nr97 homo sapien
9	10	1.2	1059	4 Q9NYG9	Q9nyg9 homo sapien
10	9	1.1	96	11 Q63I56	Q63i56 rattus norv
11	9	1.1	250	4 Q9P0Z1	Q9p0z1 homo sapien
12	9	1.1	252	10 Q94L69	Q94l69 tentilla
13	9	1.1	252	10 Q94L68	Q94l68 tentilla
14	9	1.1	360	6 Q9TTE2	Q9tte2 ovis aries
15	9	1.1	437	4 Q9ULQ7	Q9ulq7 homo sapien
16	9	1.1	575	4 Q9ULM6	Q9ulm6 homo sapien

Q9hdc7	homo sapien	626	4	Q9HDC7
Q92lu9	mus musculus	626	11	Q92LI9
Q9uk78	homo sapien	752	4	Q9UK78
Q9um57	homo sapien	799	4	Q9UM57
Q9tsp2	papio anubi	826	6	Q9TSP2
Q90206	homo sapien	839	4	Q90206
Q9ttu0	pan paniscu	839	6	Q9TTN0
Q9lnx8	arabidopsis	945	10	Q9LNX8
Q9v701	drosofila	953	5	Q9V701
Q9lkz4	glycine max	1012	10	Q9LKZ4
Q9j28	mus musculus	1271	11	Q9J28
Q9c61	arabidopsis	1784	10	Q9C61
Q9kjl0	trichodesmi	226	2	Q9KJL0
Q9nl56	pirometra	336	5	Q9NL56
Q9ved2	drosofila	344	5	Q9VED2
Q9lxu5	arabidopsis	371	10	Q9LXU5
Q94bn7	arabidopsis	371	10	Q94BN7
Q9xbw2	porphyron	428	2	Q9XBW2
Q9cxd9	mus musculus	443	11	Q9CXD9
Q96li5	homo sapien	475	4	Q96LI5
Q9zb05	lactococcus	506	2	Q9ZB05
Q9en64	spodoptera	564	12	Q9EN64
Q94143	silene lati	581	10	Q94143
Q9fhl8	arabidopsis	589	10	Q9FHL8
Q9arm3	arabidopsis	601	10	Q9ARM3
Q9arf5	capsella ru	606	10	Q9ARF5
Q9zwc6	arabidopsis	607	10	Q9ZWC6
Q94437	oryza longi	612	10	Q94437
Q86486	rinderpest	635	12	Q86486
Q9lhm5	arabidopsis	678	10	Q9LHM5
Q932w6	arabidopsis	695	10	Q932W6
Q50027	lycopersico	720	10	Q50027
Q92uh7	arabidopsi	743	10	Q92UH7
Q9zsh1	lycopersico	768	10	Q9ZSH1
Q9lfl1	arabidopsis	783	10	Q9LFL1
Q9ls79	arabidopsis	784	10	Q9LS79
Q9ma83	arabidopsis	786	10	Q9MA83
Q9y2c9	homo sapien	796	4	Q9Y2C9
Q50025	lycopersico	806	10	Q50025
Q9c815	arabidopsis	838	10	Q9C815
Q9zsh2	lycopersico	853	10	Q9ZSH2
Q50024	lycopersico	855	10	Q50024
Q9zsh3	lycopersico	855	10	Q9ZSH3
Q50026	lycopersico	862	10	Q50026
Q40235	lycopersico	863	10	Q40235
Q946d6	lycopersico	863	10	Q946D6
Q9swt8	hordeum vul	865	10	Q9SWT8
Q9zsh9	lycopersico	865	10	Q9ZSH9
Q50022	lycopersico	865	10	Q50022
Q50021	lycopersico	866	10	Q50021
Q9c905	arabidopsis	910	10	Q9C905
Q26872	trypanosoma	915	5	Q26872
Q9atq3	triticum ae	926	10	Q9ATQ3
Q9h8e3	homo sapien	928	4	Q9H8E3
Q9nyd1	homo sapien	951	4	Q9NYD1
Q9bxb1	homo sapien	951	4	Q9BXB1
Q9c6a8	arabidopsis	965	10	Q9C6A8
Q9m6a7	glycine max	987	10	Q9M6A7
Q40640	oryza sativ	1025	10	Q40640
Q9fgw1	arabidopsis	1055	10	Q9FGW1
O77690	bos taurus	1072	6	O77690
Q9lp24	arabidopsis	1120	10	Q9LP24
Q9sub9	arabidopsis	1143	10	Q9SUB9
Q9nsl	homo sapien	1176	4	Q9NSE1
Q9zvt3	arabidopsis	1187	10	Q9ZVT3
Q9p2j5	homo sapien	1212	4	Q9P2J5
Q9fhf5	arabidopsis	1214	10	Q9FHF5
Q9vjq0	drosofila	1216	5	Q9VJQ0
Q9fiz3	arabidopsis	1236	10	Q9FIT3
Q94l7	lycopersico	1271	10	Q94L7
Q9lg50	oryza sativ	1292	10	Q9LG50
Q9lrr5	arabidopsis	1424	10	Q9LRR5
Q9nk96	drosofila	1471	5	Q9NKK96

90	8	1.0	1560	10	Q9FTA6	Q9fta6 arabidopsis	163	7	0.9	250	10	Q9FWL7	Q9fwl7 oryza sativ
91	8	1.0	1590	10	Q9FTA5	Q9fta5 arabidopsis	164	7	0.9	251	12	Q9PYU2	Q9pyu2 xestia c-ni
92	7	0.9	33	5	Q25433	Q25433 molgula cit	165	7	0.9	252	10	Q94L90	Q94l90 duchesnea i
93	7	0.9	41	8	Q04378	Q04378 hordeum vul	166	7	0.9	252	10	Q94L89	Q94l89 duchesnea i
94	7	0.9	41	16	Q9A383	Q9a383 caulobacter	167	7	0.9	252	10	Q94L88	Q94l88 duchesnea i
95	7	0.9	41	16	Q92HT9	Q92ht9 rickettsia	168	7	0.9	252	10	Q94L84	Q94l84 horkelia cu
96	7	0.9	42	4	Q13497	Q13497 homo sapien	169	7	0.9	252	10	Q94L67	Q94l67 potentilla
97	7	0.9	47	11	Q9ERA2	Q9era2 microtus ar	170	7	0.9	259	2	Q86303	Q86303 streptomyce
98	7	0.9	65	10	Q08566	Q08566 catharanthu	171	7	0.9	259	2	Q05649	Q05649 streptomyce
99	7	0.9	73	2	Q9F6S9	Q9f6s9 helicobacte	172	7	0.9	260	2	Q9ADH4	Q9adh4 streptomyce
100	7	0.9	73	2	Q9F6S8	Q9f6s8 helicobacte	173	7	0.9	260	2	Q94386	Q94386 streptomyce
101	7	0.9	84	10	Q94G48	Q94g48 poncirus tr	174	7	0.9	260	11	Q97CK9	Q97ck9 thermoplasm
102	7	0.9	88	12	Q65349	Q65349 autographa	175	7	0.9	271	11	Q99L66	Q99l66 mus musculu
103	7	0.9	98	8	Q35843	Q35843 sceloporu	176	7	0.9	272	10	Q9EKS6	Q9ek6 arabidopsis
104	7	0.9	110	12	Q90623	Q90623 simian herp	177	7	0.9	273	2	Q9AEE3	Q9aee3 leptospira
105	7	0.9	110	12	Q9YMS0	Q9yts0 simian herp	178	7	0.9	274	2	Q47074	Q47074 escherichia
106	7	0.9	122	16	Q983W0	Q983w0 rhizobium l	179	7	0.9	274	2	Q9F884	Q9f884 escherichia
107	7	0.9	126	10	Q9FZF5	Q9fzf5 arabidopsis	180	7	0.9	274	2	Q9F882	Q9f882 escherichia
108	7	0.9	131	3	Q07790	Q07790 saccharomyc	181	7	0.9	274	2	Q9F877	Q9f877 escherichia
109	7	0.9	131	5	Q95T33	Q95t33 drosophila	182	7	0.9	274	2	Q9E203	Q9ez03 escherichia
110	7	0.9	134	16	Q9PFC9	Q9pfc9 xylella fas	183	7	0.9	274	2	Q9F878	Q9f878 escherichia
111	7	0.9	142	2	Q93F15	Q93f15 citrobacter	184	7	0.9	274	2	Q9F873	Q9f873 escherichia
112	7	0.9	145	16	Q9KN56	Q9kn56 vibrio chol	185	7	0.9	274	2	Q9F871	Q9f871 escherichia
113	7	0.9	148	5	Q9XUG4	Q9xug4 caenorhabdi	186	7	0.9	280	16	P74212	P74212 synechocyst
114	7	0.9	152	10	P93573	P93573 solanum tub	187	7	0.9	284	5	Q9VPC1	Q9vpc1 drosophila
115	7	0.9	155	11	Q9CV25	Q9cv25 mus musculu	188	7	0.9	284	5	Q9CV31	Q9cv31 mus musculu
116	7	0.9	156	16	Q9CG06	Q9cg06 lactococcus	189	7	0.9	287	16	Q9X099	Q9x099 thermotoga
117	7	0.9	159	2	Q9ZNJ9	Q9znj9 clostridium	190	7	0.9	290	10	Q94AF4	Q94af4 arabidopsis
118	7	0.9	159	10	Q9M636	Q9m636 glycine max	191	7	0.9	290	10	Q94AF4	Q94af4 arabidopsis
119	7	0.9	162	16	Q92HS1	Q92hs1 rickettsia	192	7	0.9	293	10	Q9LXR3	Q9lxr3 arabidopsis
120	7	0.9	163	12	Q89338	Q89338 african swi	193	7	0.9	296	16	Q92XQ7	Q92xq7 rhizobium m
121	7	0.9	166	4	Q9P1P2	Q9p1p2 homo sapien	194	7	0.9	301	10	Q9LKR34	Q9lkr34 arabidopsis
122	7	0.9	166	5	Q9N9V6	Q9n9v6 leishmania	195	7	0.9	302	17	Q59406	Q59406 pyrococcus
123	7	0.9	173	5	Q969B1	Q969b1 giardia lam	196	7	0.9	302	17	Q9V1H2	Q9v1h2 pyrococcus
124	7	0.9	179	5	Q9VCB4	Q9vcb4 drosophila	197	7	0.9	308	16	Q97KT7	Q97kt7 clostridium
125	7	0.9	183	2	Q9XBM0	Q9xbm0 acinetobact	198	7	0.9	313	12	Q91GK5	Q91gk5 epiphyas po
126	7	0.9	191	4	Q9C062	Q9c062 homo sapien	199	7	0.9	315	10	Q9LNS4	Q9lns4 arabidopsis
127	7	0.9	192	10	Q9SDR4	Q9sdr4 thuja plica	200	7	0.9	315	16	Q92A18	Q92a18 listeria in
128	7	0.9	194	2	Q52612	Q52612 borrelia bu	201	7	0.9	319	13	Q93529	Q93529 xenopus lae
129	7	0.9	194	2	Q9Z379	Q9z379 borrelia bu	202	7	0.9	321	5	Q9BHD1	Q9bhd1 leishmania
130	7	0.9	194	2	Q9ZH98	Q9zh98 borrelia bu	203	7	0.9	322	11	Q9D6Z7	Q9d6z7 mus musculu
131	7	0.9	194	10	Q65863	Q65863 pinus ellio	204	7	0.9	326	5	Q9VEK8	Q9vek8 drosophila
132	7	0.9	195	2	Q9ZH90	Q9zh90 borrelia bu	205	7	0.9	326	5	Q9U3W5	Q9u3w5 drosophila
133	7	0.9	196	16	Q9PK45	Q9pk45 chlamydia m	206	7	0.9	327	10	Q38695	Q38695 actinidia d
134	7	0.9	197	2	Q50029	Q50029 mycobacteri	207	7	0.9	329	12	Q9WHT6	Q9wh6 blue tongue
135	7	0.9	199	5	Q44230	Q44230 anthocidari	208	7	0.9	332	10	Q82438	Q82438 daucus caro
136	7	0.9	202	16	Q92B62	Q92b62 listeria in	209	7	0.9	335	7	Q77979	Q77979 rattus norv
137	7	0.9	203	16	Q92DB0	Q9x715 rhizobium m	210	7	0.9	342	2	Q87585	Q87585 campylobact
138	7	0.9	203	16	Q92DB0	Q9zdb0 rickettsia	211	7	0.9	342	2	Q9K5D3	Q9k5d3 campylobact
139	7	0.9	205	17	Q9H160	Q9h160 thermoplasm	212	7	0.9	342	16	Q9PNG1	Q9png1 campylobact
140	7	0.9	207	5	Q9W1F0	Q9w1f0 drosophila	213	7	0.9	345	16	Q9K8V2	Q9k8v2 bacillus ha
141	7	0.9	211	8	Q9G914	Q9g914 ochronomas	214	7	0.9	347	11	Q9D7D3	Q9d7d3 mus musculu
142	7	0.9	212	12	Q915F8	Q915f8 hepatitis b	215	7	0.9	348	6	Q95J38	Q95j38 macaca fasc
143	7	0.9	214	10	Q9FVZ3	Q9fvz3 oryza sativ	216	7	0.9	348	8	Q9G0N8	Q9g0n8 heterotis n
144	7	0.9	215	5	Q17578	Q17578 caenorhabdi	217	7	0.9	352	4	Q60938	Q60938 homo sapien
145	7	0.9	221	4	Q9NW37	Q9nw37 homo sapien	218	7	0.9	352	6	Q62702	Q62702 bos taurus
146	7	0.9	224	3	Q07988	Q07988 saccharomyc	219	7	0.9	353	4	Q9UJX9	Q9ujx9 homo sapien
147	7	0.9	227	16	Q9K7U1	Q9k7u1 bacillus ha	220	7	0.9	353	13	Q9DE66	Q9de66 coturnix co
148	7	0.9	230	10	Q94L83	Q94l83 horkelia cu	221	7	0.9	353	13	Q42235	Q42235 gallus gall
149	7	0.9	231	3	Q74961	Q74961 schizosacch	222	7	0.9	360	10	Q9LPL4	Q9lpl4 arabidopsis
150	7	0.9	232	16	Q9HWI7	Q9hw17 pseudomonas	223	7	0.9	361	11	Q92L05	Q92l05 mus musculu
151	7	0.9	240	10	Q9SP59	Q9sp59 populus x c	224	7	0.9	362	4	Q96T12	Q96t12 homo sapien
152	7	0.9	242	4	Q9UG10	Q9ug10 homo sapien	225	7	0.9	363	5	Q9VWV1	Q9vwv1 drosophila
153	7	0.9	244	2	Q9Z5S6	Q9z5s6 zymomonas m	226	7	0.9	364	10	Q9C879	Q9c879 arabidopsis
154	7	0.9	245	2	Q9F809	Q9f8u9 streptomyce	227	7	0.9	364	16	Q9CBC2	Q9cbc2 mycobacteri
155	7	0.9	245	8	Q95FD5	Q95fd5 lunularia c	228	7	0.9	370	4	Q9NQA7	Q9nqa7 homo sapien
156	7	0.9	246	10	Q94L47	Q94l47 spiraea can	229	7	0.9	371	10	Q9FHH3	Q9fhh3 arabidopsis
157	7	0.9	248	16	Q9HWX9	Q9hwx9 pseudomonas	230	7	0.9	371	10	Q9AUN5	Q9aun5 oryza sativ
158	7	0.9	249	2	Q9L575	Q9l575 streptococ	231	7	0.9	371	13	Q9W676	Q9w676 gallus gall
159	7	0.9	249	10	Q94L85	Q94l85 holodiscus	232	7	0.9	373	4	Q9P035	Q9p035 homo sapien
160	7	0.9	249	10	Q94L46	Q94l46 spiraea den	233	7	0.9	376	10	Q9LXP9	Q9lpx9 arabidopsis
161	7	0.9	249	10	Q94L45	Q94l45 spiraea den	234	7	0.9	377	3	Q00012	Q00012 aspergillus
162	7	0.9	250	5	Q02212	Q02212 caenorhabdi	235	7	0.9	378	6	Q9XS10	Q9xs10 bos indicus

236	7	0.9	380	11	Q61076	Q61076 mus musculus	309	7	0.9	545	10	O80957	O80957 arabidopsis
237	7	0.9	381	16	Q98H70	Q98H70 rhizobium l	310	7	0.9	549	10	Q9SVW8	Q9SVW8 arabidopsis
238	7	0.9	387	5	Q23459	caenorhabdi	311	7	0.9	550	5	Q9VJN8	Q9VJN8 drosophila
239	7	0.9	389	10	O22313	lycopersico	312	7	0.9	555	16	Q97K3	Q97K3 streptococ
240	7	0.9	389	5	Q95U18	lycopersico	313	7	0.9	555	3	Q9HDP8	Q9HDP8 schizosacch
241	7	0.9	389	10	Q9FNB8	arabidopsis	314	7	0.9	557	10	Q94C39	Q94C39 arabidopsis
242	7	0.9	389	2	Q9LAZ3	lycopersico	315	7	0.9	562	5	Q21164	Q21164 caenorhabdi
243	7	0.9	394	5	Q9VXF3	drosophila	316	7	0.9	562	10	Q9M7W9	Q9M7W9 arabidopsis
244	7	0.9	395	2	Q9LAY2	streptococ	317	7	0.9	564	2	Q9S468	Q9S468 bacillus st
245	7	0.9	395	10	O48581	arabidopsis	318	7	0.9	566	16	P72770	P72770 synechocyst
246	7	0.9	396	4	Q9NQS5	homo sapien	319	7	0.9	568	9	Q37989	Q37989 bacterioph
247	7	0.9	400	2	Q9XDP7	acinetobact	320	7	0.9	570	10	Q9FK65	Q9FK65 arabidopsis
248	7	0.9	402	10	O82202	arabidopsis	321	7	0.9	572	17	Q980B9	Q980B9 sulfolobus
249	7	0.9	408	2	Q9LAY0	streptococ	322	7	0.9	577	16	Q9KE87	Q9KE87 bacillus ha
250	7	0.9	410	16	O84710	chlamydia t	323	7	0.9	584	10	O49750	O49750 arabidopsis
251	7	0.9	415	10	Q9M0F7	arabidopsis	324	7	0.9	584	10	O49751	O49751 arabidopsis
252	7	0.9	417	2	Q9R6S6	synechococ	325	7	0.9	589	3	O94047	O94047 candida alb
253	7	0.9	417	5	O01618	caenorhabdi	326	7	0.9	593	8	Q9G8P7	Q9G8P7 naegleria g
254	7	0.9	417	16	P96441	rhizobium m	327	7	0.9	593	10	Q94IM4	Q94IM4 hordeum vul
255	7	0.9	418	16	P76526	escherichia	328	7	0.9	594	10	Q9LUQ2	Q9LUQ2 arabidopsis
256	7	0.9	419	10	Q9M1B6	arabidopsis	329	7	0.9	604	10	Q9S185	Q9S185 arabidopsis
257	7	0.9	421	5	Q9XVS8	caenorhabdi	330	7	0.9	608	4	Q9Y627	Q9Y627 homo sapien
258	7	0.9	425	10	O9FK66	arabidopsis	331	7	0.9	610	4	Q96GJ2	Q96GJ2 homo sapien
259	7	0.9	426	2	Q9LAY5	streptococ	332	7	0.9	611	12	Q9YW82	Q9YW82 melanoplus
260	7	0.9	426	11	Q9D655	mus musculus	333	7	0.9	613	10	Q940E8	Q940E8 zea mays (m
261	7	0.9	429	2	P70750	acetobacter	334	7	0.9	617	10	Q9FJF1	Q9FJF1 arabidopsis
262	7	0.9	429	12	Q9YVK3	melanoplus	335	7	0.9	619	10	Q9XGG1	Q9XGG1 sorghum bic
263	7	0.9	430	8	Q9XMS1	tetrahymena	336	7	0.9	623	16	Q9PJ67	Q9PJ67 campylobact
264	7	0.9	438	10	O9LEGO	lycopersico	337	7	0.9	624	11	O54824	O54824 mus musculu
265	7	0.9	438	16	Q9X1L8	thermotoga	338	7	0.9	624	11	O70236	O70236 mus musculu
266	7	0.9	440	16	Q9CIL4	lactococcus	339	7	0.9	624	12	Q9YV11	Q9YV11 melanoplus
267	7	0.9	443	2	Q9X365	bacillus an	340	7	0.9	625	2	O30704	O30704 mycoplasma
268	7	0.9	446	10	O9F6N4	arabidopsis	341	7	0.9	627	3	O30643	O30643 mycoplasma
269	7	0.9	447	2	O52382	ralstonia s	342	7	0.9	627	5	Q9N9B1	Q9N9B1 leishmania
270	7	0.9	447	2	O9L5D2	comamonas t	343	7	0.9	627	6	O62677	O62677 saimiri sci
271	7	0.9	447	2	P95564	pseudomonas	344	7	0.9	627	16	P72862	P72862 synechocyst
272	7	0.9	448	16	Q9PLJ3	chlamydia m	345	7	0.9	630	1	O93675	O93675 sulfolobus
273	7	0.9	451	2	O45695	burkholderi	346	7	0.9	630	5	O16366	O16366 caenorhabdi
274	7	0.9	453	5	O23615	caenorhabdi	347	7	0.9	630	6	O62675	O62675 macaca mula
275	7	0.9	453	5	O9VZ14	drosophila	348	7	0.9	630	6	O62676	O62676 macaca fasc
276	7	0.9	454	5	O9VXK6	lycopersico	349	7	0.9	631	4	O9UME6	O9UME6 homo sapien
277	7	0.9	456	11	O924P4	mus musculus	350	7	0.9	631	6	O62666	O62666 pan troglod
278	7	0.9	456	11	O924P4	mus musculus	351	7	0.9	631	6	O62674	O62674 cercopithe
279	7	0.9	456	11	O91V17	caenorhabdi	352	7	0.9	632	6	O62678	O62678 aotus trivi
280	7	0.9	457	5	Q9XU95	caenorhabdi	353	7	0.9	638	5	Q17989	Q17989 caenorhabdi
281	7	0.9	459	5	Q9V4W7	drosophila	354	7	0.9	638	10	O9LFS4	O9LFS4 arabidopsis
282	7	0.9	464	10	O39619	catharanthu	355	7	0.9	641	10	O9SKB2	O9SKB2 arabidopsis
283	7	0.9	474	4	O965M8	homo sapien	356	7	0.9	641	10	Q93240	Q93240 arabidopsis
284	7	0.9	476	10	Q9LNV8	arabidopsis	357	7	0.9	645	2	Q9X6C6	Q9X6C6 thermus bro
285	7	0.9	480	10	O48705	arabidopsis	358	7	0.9	650	10	O9SR24	O9SR24 arabidopsis
286	7	0.9	490	4	Q9H5N5	homo sapien	359	7	0.9	654	10	O92V25	O92V25 arabidopsis
287	7	0.9	490	10	Q9FSS8	oryza sativ	360	7	0.9	659	2	O59289	O59289 pedobacter
288	7	0.9	492	5	Q917K6	drosophila	361	7	0.9	660	4	O43155	O43155 homo sapien
289	7	0.9	492	13	Q98T96	brachydanio	362	7	0.9	661	4	Q99467	Q99467 homo sapien
290	7	0.9	493	10	O9LR45	arabidopsis	363	7	0.9	662	10	Q9SN49	Q9SN49 arabidopsis
291	7	0.9	494	17	O58866	pyrococcus	364	7	0.9	662	10	Q9LQN6	Q9LQN6 arabidopsis
292	7	0.9	500	5	Q9N9H9	leishmania	365	7	0.9	662	10	Q9S9Q9	Q9S9Q9 arabidopsis
293	7	0.9	504	11	Q99M43	mus musculus	366	7	0.9	665	16	Q92E08	Q92E08 listeria in
294	7	0.9	505	5	Q9V9Y0	drosophila	367	7	0.9	666	10	O9FK63	O9FK63 arabidopsis
295	7	0.9	505	10	O9FFJ3	arabidopsis	368	7	0.9	674	10	Q9LZ86	Q9LZ86 arabidopsis
296	7	0.9	509	16	O07622	bacillus su	369	7	0.9	677	5	Q9VEK6	Q9VEK6 drosophila
297	7	0.9	512	10	Q9FW47	arabidopsis	370	7	0.9	677	5	Q9W266	Q9W266 drosophila
298	7	0.9	513	11	O62661	rattus norv	371	7	0.9	677	10	Q9LNV9	Q9LNV9 arabidopsis
299	7	0.9	514	4	O60381	homo sapien	372	7	0.9	684	5	O9BI68	O9BI68 caenorhabdi
300	7	0.9	514	4	O14790	homo sapien	373	7	0.9	685	10	O22123	O22123 arabidopsis
301	7	0.9	514	4	O96AJ2	homo sapien	374	7	0.9	685	10	Q9M0L9	Q9M0L9 arabidopsis
302	7	0.9	524	10	Q9LDE9	oryza sativ	375	7	0.9	685	10	Q9FY46	Q9FY46 arabidopsis
303	7	0.9	527	10	Q9NAM4	arabidopsis	376	7	0.9	685	10	Q9C9N5	Q9C9N5 arabidopsis
304	7	0.9	531	10	O65536	arabidopsis	377	7	0.9	687	16	Q92NF3	Q92NF3 rhizobium m
305	7	0.9	532	10	O96226	adiantum ca	378	7	0.9	688	10	O9C819	O9C819 arabidopsis
306	7	0.9	537	10	O9C769	arabidopsis	379	7	0.9	691	10	Q93YK5	Q93YK5 brassica na
307	7	0.9	540	5	Q9N884	drosophila	380	7	0.9	696	4	Q9UFC0	Q9UFC0 homo sapien
308	7	0.9	540	10	P93666	heliathanus	381	7	0.9	697	3	O74504	O74504 schizosacch

382	7	0.9	702	5	Q93539	Q93539 caenorhabdi	455	7	0.9	883	10	Q9LRW9	Q91rw9 arabisdopsis
383	7	0.9	702	10	Q9FRI1	Q9fril1 arabisdopsis	456	7	0.9	883	10	Q9FZB7	Q9fbz7 arabisdopsis
384	7	0.9	707	12	Q9YV15	Q9yv15 melanoplus	457	7	0.9	884	10	Q91069	Q91069 arabisdopsis
385	7	0.9	708	4	Q9H075	Q9h075 homo sapien	458	7	0.9	889	10	Q9SGU1	Q9sgul arabisdopsis
386	7	0.9	709	11	Q9J023	Q9jj23 mus musculus	459	7	0.9	890	10	Q98849	Q98849 arabisdopsis
387	7	0.9	711	10	Q9LJW7	Q9ljw7 arabisdopsis	460	7	0.9	890	10	Q9LJS0	Q91js0 arabisdopsis
388	7	0.9	718	13	Q9J3675	Q9j3675 xenopus lae	461	7	0.9	890	10	Q93ZT8	Q93zt8 arabisdopsis
389	7	0.9	721	4	Q9NUV1	Q9nuv1 homo sapien	462	7	0.9	892	10	Q92484	Q92484 arabisdopsis
390	7	0.9	725	10	Q9L106	Q9l106 zea mays (m	463	7	0.9	893	10	Q9SXD4	Q9sxd4 arabisdopsis
391	7	0.9	728	10	Q9M9S4	Q9m9s4 arabisdopsis	464	7	0.9	896	10	Q9W9B0	Q9w9b0 arabisdopsis
392	7	0.9	732	4	Q9P0V2	Q9p0v2 homo sapien	465	7	0.9	899	10	Q9LIG2	Q9llg2 arabisdopsis
393	7	0.9	732	4	Q9BQ33	Q9bq33 homo sapien	466	7	0.9	901	10	Q9C6G6	Q9c6g6 arabisdopsis
394	7	0.9	732	6	Q9TS46	Q9ts46 sus scrofa	467	7	0.9	904	4	Q15455	Q15455 homo sapien
395	7	0.9	738	13	Q90245	Q90245 gallus gall	468	7	0.9	904	5	Q9VMJ4	Q9vmj4 drosophila
396	7	0.9	739	10	Q90495	Q90495 nicotiana t	469	7	0.9	904	10	Q9SCN7	Q9scn7 arabisdopsis
397	7	0.9	740	10	Q94325	Q94325 arabisdopsis	470	7	0.9	905	10	Q64757	Q64757 arabisdopsis
398	7	0.9	741	10	Q9SVN2	Q9svn2 arabisdopsis	471	7	0.9	907	10	Q9FW49	Q9fw49 arabisdopsis
399	7	0.9	744	10	Q9SHU6	Q9shu6 arabisdopsis	472	7	0.9	915	5	Q9W267	Q9w267 drosophila
400	7	0.9	744	10	Q9C8I8	Q9c8i8 arabisdopsis	473	7	0.9	928	4	Q9BYD7	Q9byd7 homo sapien
401	7	0.9	751	10	Q9FGQ5	Q9fgq5 arabisdopsis	474	7	0.9	933	5	Q9BJD5	Q9bjd5 strongyloce
402	7	0.9	751	10	Q94DM7	Q94dm7 arabisdopsis	475	7	0.9	949	8	Q98RX5	Q98rx5 guilladia
403	7	0.9	753	4	Q969L7	Q969l7 homo sapien	476	7	0.9	949	16	Q92HW8	Q92hw8 rickettsia
404	7	0.9	756	10	Q9SHI4	Q9shi4 arabisdopsis	477	7	0.9	950	13	Q90Z44	Q90z44 gallus gall
405	7	0.9	756	16	Q9RZS5	Q9rzs5 delinococcus	478	7	0.9	950	16	Q51770	Q51770 borrelia bu
406	7	0.9	757	3	P87109	P87109 schizosacch	479	7	0.9	951	10	Q9C6A7	Q9c6a7 arabisdopsis
407	7	0.9	758	5	Q22203	Q22203 caenorhabdi	480	7	0.9	951	11	Q922H4	Q922h4 rattus norv
408	7	0.9	764	16	Q67347	Q67347 aquifex aeo	481	7	0.9	956	10	Q9AR02	Q9ar02 hordeum vul
409	7	0.9	765	10	Q9L455	Q9l455 arabisdopsis	482	7	0.9	957	10	Q9S9U3	Q9s9u3 arabisdopsis
410	7	0.9	768	5	Q9VCR4	Q9vcr4 drosophila	483	7	0.9	958	5	Q969B4	Q969b4 giardia lam
411	7	0.9	770	10	Q9FL15	Q9fl15 arabisdopsis	484	7	0.9	958	10	Q9FE39	Q9fe39 hordeum vul
412	7	0.9	771	10	Q48851	Q48851 arabisdopsis	485	7	0.9	960	3	Q9USQ1	Q9usq1 schizosacch
413	7	0.9	773	10	Q98RM5	Q98rm5 guilladia	486	7	0.9	960	5	Q9NCP5	Q9ncp5 giardia lam
414	7	0.9	783	13	Q90XG2	Q90xg2 gallus gall	487	7	0.9	964	10	Q9LY03	Q9ly03 arabisdopsis
415	7	0.9	784	10	Q9C9H6	Q9c9h6 arabisdopsis	488	7	0.9	976	10	Q42371	Q42371 arabisdopsis
416	7	0.9	787	3	Q96X16	Q96x16 pichia past	489	7	0.9	988	10	Q9SN81	Q9sn81 arabisdopsis
417	7	0.9	791	5	Q94170	Q94170 caenorhabdi	490	7	0.9	991	10	Q94C44	Q94c44 chlamydomon
418	7	0.9	793	5	Q9VNH0	Q9vnh0 drosophila	491	7	0.9	992	10	Q65440	Q65440 arabisdopsis
419	7	0.9	796	11	Q9WVC1	Q9wvc1 rattus norv	492	7	0.9	993	4	Q43352	Q43352 homo sapien
420	7	0.9	799	4	Q9P0U2	Q9p0u2 homo sapien	493	7	0.9	999	10	Q9S711	Q9s711 oryza sativ
421	7	0.9	807	4	Q9Y2R2	Q9y2r2 homo sapien	494	7	0.9	1002	10	Q9M2Z1	Q9m2z1 arabisdopsis
422	7	0.9	808	4	Q95063	Q95063 homo sapien	495	7	0.9	1007	10	Q9ZVD4	Q9zvd4 arabisdopsis
423	7	0.9	808	10	Q9ZWT4	Q9zwt4 ipomoea pur	496	7	0.9	1008	10	Q9LXZ6	Q9lxz6 glycine max
424	7	0.9	809	1	Q9P9C4	Q9p9c4 uncultured	497	7	0.9	1008	10	Q943S4	Q943s4 oryza sativ
425	7	0.9	811	4	Q9P2B1	Q9p2b1 homo sapien	498	7	0.9	1019	10	Q9C699	Q9c699 arabisdopsis
426	7	0.9	811	4	Q9BXR5	Q9bxr5 homo sapien	499	7	0.9	1020	10	Q22579	Q22579 arabisdopsis
427	7	0.9	812	5	Q9VFY9	Q9vfy9 drosophila	500	7	0.9	1021	5	Q9V430	Q9v430 drosophila
428	7	0.9	812	10	Q9FWT0	Q9fmt0 arabisdopsis	501	7	0.9	1021	10	Q9FXF2	Q9fxf2 arabisdopsis
429	7	0.9	818	10	Q9SLI6	Q9sli6 arabisdopsis	502	7	0.9	1022	10	Q9LCK8	Q9lgk8 oryza sativ
430	7	0.9	820	11	Q9P5S7	Q9d5s7 mus musculus	503	7	0.9	1025	11	Q9Z166	Q9z166 mus musculus
431	7	0.9	823	5	Q23141	Q23141 caenorhabdi	504	7	0.9	1027	10	Q9LC26	Q9lc26 arabisdopsis
432	7	0.9	825	4	Q96M69	Q96m69 homo sapien	505	7	0.9	1032	11	Q9EQU3	Q9equ3 mus musculus
433	7	0.9	835	10	Q9LS80	Q9ls80 arabisdopsis	506	7	0.9	1032	11	Q99MQ8	Q99mq8 mus musculus
434	7	0.9	835	11	Q9QX05	Q9qxo5 rattus norv	507	7	0.9	1032	11	Q99MF2	Q99mf2 mus musculus
435	7	0.9	838	11	Q9WQW5	Q9wxw5 cricetus	508	7	0.9	1036	10	Q9FN37	Q9fn37 arabisdopsis
436	7	0.9	839	9	Q64076	Q64076 bacterioph	509	7	0.9	1038	12	Q918V1	Q918v1 bovine parv
437	7	0.9	839	16	Q31945	Q31945 bacillus su	510	7	0.9	1038	5	Q24007	Q24007 drosophila
438	7	0.9	845	10	Q94879	Q94879 lycopersico	511	7	0.9	1051	5	Q95PA9	Q95pa9 drosophila
439	7	0.9	845	10	Q9C9H7	Q9c9h7 arabisdopsis	512	7	0.9	1060	5	Q9I7Q3	Q9i7q3 drosophila
440	7	0.9	848	2	Q9RAT4	Q9rat4 lactococcus	513	7	0.9	1066	5	Q18902	Q18902 caenorhabdi
441	7	0.9	848	16	Q9CHQ5	Q9chq5 lactococcus	514	7	0.9	1072	11	Q9DBT7	Q9dbt7 mus musculus
442	7	0.9	852	16	Q9WZG5	Q9wzg5 thermotoga	515	7	0.9	1077	5	Q9NJG7	Q9njg7 drosophila
443	7	0.9	857	10	Q94125	Q94125 oryza sativ	516	7	0.9	1078	16	Q9A3J0	Q9a3j0 caulobacter
444	7	0.9	858	4	Q60346	Q60346 homo sapien	517	7	0.9	1079	10	Q9CA77	Q9ca77 arabisdopsis
445	7	0.9	859	11	Q9JILF7	Q9jilf7 mus musculus	518	7	0.9	1086	10	Q9SGU0	Q9sgu0 arabisdopsis
446	7	0.9	861	10	Q9SLS3	Q9sls3 nicotiana t	519	7	0.9	1100	5	Q24622	Q24622 drosophila
447	7	0.9	862	10	Q50020	Q50020 lycopersico	520	7	0.9	1108	10	Q9ARB2	Q9arb2 linum usita
448	7	0.9	865	10	Q50028	Q50028 lycopersico	521	7	0.9	1108	10	Q9ARB0	Q9arb0 linum usita
449	7	0.9	865	10	Q50023	Q50023 lycopersico	522	7	0.9	1110	10	Q94LN2	Q94ln2 oryza sativ
450	7	0.9	868	10	Q9FN94	Q9fn94 arabisdopsis	523	7	0.9	1116	5	Q9XYW1	Q9xyw1 homarus ame
451	7	0.9	868	10	Q9M9X0	Q9m9x0 arabisdopsis	524	7	0.9	1120	10	Q9LGB1	Q9lgb1 oryza sativ
452	7	0.9	871	10	Q9FZB8	Q9fzb8 arabisdopsis	525	7	0.9	1121	2	Q9AKN9	Q9akn9 rickettsia
453	7	0.9	875	10	Q9FZB4	Q9fzb4 arabisdopsis	526	7	0.9	1121	3	Q12734	Q12734 saccharomyc
454	7	0.9	879	10	Q9FKC2	Q9fkC2 arabisdopsis	527	7	0.9	1122	2	Q9AKI6	Q9aki6 rickettsia

528	7	0.9	1122	16	Q92H58	Q92H58 rickettsia	601	7	0.9	2204	12	Q90341	O90341 newcastle d
529	7	0.9	1123	5	Q9W3W0	Q9W3W0 drosophila	602	7	0.9	2204	12	Q9DLD3	Q9DLD3 newcastle d
530	7	0.9	1123	10	Q9SWSL9	Q9SWSL9 arabisdopsis	603	7	0.9	2204	12	Q9WMH6	Q9WMH6 newcastle d
531	7	0.9	1125	5	Q93203	Q93203 caenorhabdi	604	7	0.9	2204	12	Q91AH6	Q91AH6 newcastle d
532	7	0.9	1125	10	Q94DJ2	Q94DJ2 oryza sativ	605	7	0.9	2307	2	Q9AG79	Q9AG79 streptomyce
533	7	0.9	1145	10	Q9FEU2	Q9FEU2 pinus sylve	606	7	0.9	2342	12	Q9G5980	Q9G5980 cherry capi
534	7	0.9	1156	5	Q967X9	Q967X9 tribolium c	607	7	0.9	2343	6	Q962730	Q962730 canis famil
535	7	0.9	1166	10	Q9ZWC8	Q9ZWC8 arabisdopsis	608	7	0.9	2343	6	Q918806	Q918806 canis famil
536	7	0.9	1166	10	Q9ARF3	Q9ARF3 capsella ru	609	7	0.9	2437	6	Q95MJ3	Q95MJ3 oryctolagus
537	7	0.9	1173	10	Q9FL28	Q9FL28 arabisdopsis	610	7	0.9	2454	3	Q9UVP2	Q9UVP2 emericella
538	7	0.9	1179	12	Q94919	Q94919 pepper ring	611	7	0.9	2454	3	Q9UV56	Q9UV56 emericella
539	7	0.9	1187	10	Q93VG9	Q93VG9 oryza sativ	612	7	0.9	2473	11	Q9QZ84	Q9QZ84 mus musculus
540	7	0.9	1192	10	Q9LYN8	Q9LYN8 arabisdopsis	613	7	0.9	3036	5	Q917575	Q917575 caenorhabdi
541	7	0.9	1196	10	Q92476	Q92476 arabisdopsis	614	7	0.9	3246	5	Q9GY61	Q9GY61 leishmania
542	7	0.9	1197	4	Q95347	Q95347 homo sapien	615	7	0.9	3477	11	Q99MH8	Q99MH8 mus musculus
543	7	0.9	1202	17	Q97B44	Q97B44 thermoplasma	616	7	0.9	4083	3	Q9C1M7	Q9C1M7 ashbya goss
544	7	0.9	1209	5	Q94525	Q94525 drosophila	617	7	0.9	4099	10	Q9C726	Q9C726 arabisdopsis
545	7	0.9	1237	16	Q9XCG8	Q9XCG8 listeria in	618	7	0.9	4199	16	P74440	P74440 synecocyst
546	7	0.9	1266	10	Q9XET3	Q9XET3 lycopersico	619	7	0.9	4202	11	Q91XP9	Q91XP9 mus musculus
547	7	0.9	1274	5	Q90985	Q90985 dictyosteli	620	7	0.9	4202	11	Q91V63	Q91V63 mus musculus
548	7	0.9	1291	11	Q61812	Q61812 mus musculus	621	7	0.9	4293	11	O08852	O08852 mus musculus
549	7	0.9	1297	10	Q945S6	Q945S6 lycopersico	622	7	0.9	4307	5	Q19319	Q19319 caenorhabdi
550	7	0.9	1301	10	Q92350	Q92350 arabisdopsis	623	7	0.9	4345	5	Q9VLA0	Q9VLA0 drosophila
551	7	0.9	1304	10	Q9SUK4	Q9SUK4 arabisdopsis	624	7	0.9	4370	4	Q9H3V5	Q9H3V5 homo sapien
552	7	0.9	1317	10	Q923535	Q923535 arabisdopsis	625	7	0.9	4523	4	Q96D75	Q96D75 homo sapien
553	7	0.9	1322	11	Q9QZP6	Q9QZP6 mus musculus	626	7	0.9	4717	3	Q94248	Q94248 schizosacch
554	7	0.9	1326	5	Q9W4N4	Q9W4N4 drosophila	627	7	0.9	4731	11	Q91XQ0	Q91XQ0 mus musculus
555	7	0.9	1343	11	Q64730	Q64730 mus musculus	628	7	0.9	4731	11	Q91XP8	Q91XP8 mus musculus
556	7	0.9	1344	10	Q9SM94	Q9SM94 oryza sativ	629	6	0.7	14	11	O70599	O70599 rattus norv
557	7	0.9	1348	5	Q9VAD1	Q9VAD1 drosophila	630	6	0.7	27	11	Q9C5C1	Q9C5C1 mus musculus
558	7	0.9	1354	11	Q9EPW8	Q9EPW8 mus musculus	631	6	0.7	28	4	Q16326	Q16326 homo sapien
559	7	0.9	1361	10	Q04264	Q04264 arabisdopsis	632	6	0.7	33	2	O87520	O87520 escherichia
560	7	0.9	1369	5	Q9V4R0	Q9V4R0 drosophila	633	6	0.7	39	16	Q932N4	Q932N4 staphylococ
561	7	0.9	1379	5	Q9V894	Q9V894 drosophila	634	6	0.7	41	12	O90R80	O90R80 hepatitis b
562	7	0.9	1392	5	Q95TQ7	Q95TQ7 drosophila	635	6	0.7	41	16	Q92M65	Q92M65 rhizobium m
563	7	0.9	1420	10	Q9XE13	Q9XE13 oryza sativ	636	6	0.7	42	12	Q9E0V1	Q9E0V1 hepatitis b
564	7	0.9	1436	3	Q07527	Q07527 saccharomyc	637	6	0.7	44	9	Q37937	Q37937 lactococcus
565	7	0.9	1440	5	Q20204	Q20204 caenorhabdi	638	6	0.7	45	2	Q49069	Q49069 mycoplasma
566	7	0.9	1443	10	Q9FLY5	Q9FLY5 arabisdopsis	639	6	0.7	48	5	Q9VNP1	Q9VNP1 drosophila
567	7	0.9	1448	16	Q9HYW9	Q9HYW9 pseudomonas	640	6	0.7	49	11	Q99011	Q99011 rattus norv
568	7	0.9	1454	3	Q06164	Q06164 saccharomyc	641	6	0.7	49	16	Q97S86	Q97S86 streptococc
569	7	0.9	1460	4	Q9C0A1	Q9C0A1 homo sapien	642	6	0.7	52	4	Q96QR9	Q96QR9 homo sapien
570	7	0.9	1495	11	P70587	P70587 rattus norv	643	6	0.7	53	10	Q9XG02	Q9XG02 arabisdopsis
571	7	0.9	1504	4	Q9UES6	Q9UES6 homo sapien	644	6	0.7	54	16	Q98B39	Q98B39 rhizobium l
572	7	0.9	1513	5	Q9VFP0	Q9VFP0 drosophila	645	6	0.7	55	11	Q9WUN4	Q9WUN4 rattus norv
573	7	0.9	1521	4	Q95710	Q95710 homo sapien	646	6	0.7	56	6	P79283	P79283 sus scrofa
574	7	0.9	1521	11	Q9R1B9	Q9R1B9 mus musculus	647	6	0.7	56	12	Q9WP60	Q9WP60 hepatitis b
575	7	0.9	1525	4	Q9Y5Q7	Q9Y5Q7 homo sapien	648	6	0.7	56	17	Q97YS2	Q97YS2 sulfolobus
576	7	0.9	1528	4	Q9Y2I1	Q9Y2I1 homo sapien	649	6	0.7	57	16	Q98LP1	Q98LP1 rhizobium l
577	7	0.9	1529	4	Q94813	Q94813 homo sapien	650	6	0.7	59	10	Q9SPW8	Q9SPW8 picea glauc
578	7	0.9	1530	11	Q9WUG5	Q9WUG5 rattus norv	651	6	0.7	59	12	Q9PX82	Q9PX82 hepatitis b
579	7	0.9	1531	11	O88279	O88279 rattus norv	652	6	0.7	60	11	O99019	O99019 rattus norv
580	7	0.9	1531	11	Q9WVB5	Q9WVB5 mus musculus	653	6	0.7	61	1	Q9C4W7	Q9C4W7 sulfolobus
581	7	0.9	1534	4	Q75093	Q75093 homo sapien	654	6	0.7	62	10	Q9AXV6	Q9AXV6 brassica na
582	7	0.9	1537	4	Q96NW7	Q96NW7 homo sapien	655	6	0.7	63	10	Q9LFW6	Q9LFW6 arabisdopsis
583	7	0.9	1600	10	Q9SM84	Q9SM84 oryza sativ	656	6	0.7	63	17	Q975W4	Q975W4 sulfolobus
584	7	0.9	1603	4	Q15737	Q15737 homo sapien	657	6	0.7	64	2	Q9R7D8	Q9R7D8 mycobacteri
585	7	0.9	1615	5	Q95X46	Q95X46 caenorhabdi	658	6	0.7	64	5	O02070	O02070 caenorhabdi
586	7	0.9	1661	10	Q9SOW9	Q9SOW9 arabisdopsis	659	6	0.7	64	12	Q9QRR2	Q9QRR2 hepatitis b
587	7	0.9	1696	11	Q9WTR8	Q9WTR8 rattus norv	660	6	0.7	64	16	Q92H47	Q92H47 rickettsia
588	7	0.9	1698	12	Q84918	Q84918 pepper ring	661	6	0.7	65	12	O69376	O69376 hepatitis b
589	7	0.9	1724	5	Q96960	Q96960 drosophila	662	6	0.7	65	12	O69377	O69377 hepatitis b
590	7	0.9	1726	11	O62383	O62383 mus musculus	663	6	0.7	65	12	O69378	O69378 hepatitis b
591	7	0.9	1789	10	Q9LNB5	Q9LNB5 arabisdopsis	664	6	0.7	65	12	O69379	O69379 hepatitis b
592	7	0.9	1810	5	Q9V483	Q9V483 drosophila	665	6	0.7	65	12	O9J3T9	O9J3T9 hepatitis b
593	7	0.9	1813	10	Q9ZSC9	Q9ZSC9 lactuca sat	666	6	0.7	65	12	O9J3T8	O9J3T8 hepatitis b
594	7	0.9	1968	12	Q89548	Q89548 potato viru	667	6	0.7	65	12	O9J3T7	O9J3T7 hepatitis b
595	7	0.9	2018	5	Q20487	Q20487 caenorhabdi	668	6	0.7	65	12	O9J3T6	O9J3T6 hepatitis b
596	7	0.9	2049	10	Q9M378	Q9M378 arabisdopsis	669	6	0.7	65	12	O9J3T5	O9J3T5 hepatitis b
597	7	0.9	2057	10	Q9AW36	Q9AW36 guillardia	670	6	0.7	65	12	O9J3T4	O9J3T4 hepatitis b
598	7	0.9	2062	10	Q9C7J2	Q9C7J2 arabisdopsis	671	6	0.7	65	12	Q91F62	Q91F62 chilo iride
599	7	0.9	2155	3	Q9HEP4	Q9HEP4 blumeria gr	672	6	0.7	66	12	O9WJE7	O9WJE7 hepatitis b
600	7	0.9	2198	5	Q18990	Q18990 caenorhabdi	673	6	0.7	66	12	O91SZ7	O91SZ7 hepatitis b

674	6	0.7	66	12	Q91A86	Q91A86 hepatitis b	747	6	0.7	86	10	Q944F0	Q944F0 oryza sativ
675	6	0.7	66	12	Q91A84	Q91A84 hepatitis b	748	6	0.7	86	16	Q98P96	Q98P96 rhizobium l
676	6	0.7	66	12	Q91A82	Q91A82 hepatitis b	749	6	0.7	86	16	Q99VD3	Q99VD3 staphylococ
677	6	0.7	66	12	Q91A80	Q91A80 hepatitis b	750	6	0.7	87	2	Q93HK7	Q93HK7 streptomyce
678	6	0.7	66	12	Q91A76	Q91A76 hepatitis b	751	6	0.7	87	12	Q9E9A2	Q9E9A2 hepatitis b
679	6	0.7	66	12	Q91A74	Q91A74 hepatitis b	752	6	0.7	88	6	O62851	O62851 oviss aries
680	6	0.7	66	12	Q91A72	Q91A72 hepatitis b	753	6	0.7	88	10	Q9LN99	Q9LN99 arabisopsis
681	6	0.7	66	12	Q91A70	Q91A70 hepatitis b	754	6	0.7	88	13	Q9DER3	Q9DER3 gallus galli
682	6	0.7	66	12	Q91A68	Q91A68 hepatitis b	755	6	0.7	88	16	Q97F86	Q97F86 clostridium
683	6	0.7	66	12	Q91A66	Q91A66 hepatitis b	756	6	0.7	89	7	Q9GJ13	Q9GJ13 homo sapien
684	6	0.7	66	12	Q91A60	Q91A60 hepatitis b	757	6	0.7	89	10	Q9SPX0	Q9SPX0 picea abies
685	6	0.7	66	12	Q91A58	Q91A58 hepatitis b	758	6	0.7	89	12	Q914J3	Q914J3 sulfolobus
686	6	0.7	66	12	Q91A56	Q91A56 hepatitis b	759	6	0.7	89	16	Q9HV58	Q9HV58 pseudomonas
687	6	0.7	66	12	Q91A54	Q91A54 hepatitis b	760	6	0.7	89	16	Q932P3	Q932P3 staphylococ
688	6	0.7	66	12	Q91A53	Q91A53 hepatitis b	761	6	0.7	90	2	Q93BD8	Q93BD8 salmonella
689	6	0.7	66	12	Q91A52	Q91A52 hepatitis b	762	6	0.7	90	12	O39898	O39898 hepatitis c
690	6	0.7	66	12	Q91A51	Q91A51 hepatitis b	763	6	0.7	91	2	Q93BC9	Q93BC9 salmonella
691	6	0.7	66	12	Q91A50	Q91A50 hepatitis b	764	6	0.7	91	8	Q9B968	Q9B968 ceratosen
692	6	0.7	66	12	Q91A49	Q91A49 hepatitis b	765	6	0.7	91	12	Q67836	Q67836 hepatitis b
693	6	0.7	66	12	Q91A48	Q91A48 hepatitis b	766	6	0.7	92	2	Q93BA7	Q93BA7 salmonella
694	6	0.7	66	12	Q91A47	Q91A47 hepatitis b	767	6	0.7	92	2	Q933B6	Q933B6 salmonella
695	6	0.7	67	2	Q931N4	Q931N4 salmonella	768	6	0.7	93	2	Q93B98	Q93B98 salmonella
696	6	0.7	67	3	O14391	O14391 schizosacch	769	6	0.7	94	11	Q05702	Q05702 rattus ratt
697	6	0.7	67	5	Q9V971	Q9V971 drosophila	770	6	0.7	94	12	Q9WMW7	Q9WMW7 hepatitis b
698	6	0.7	68	2	Q9S1S1	Q9S1S1 streptomyce	771	6	0.7	94	12	Q9DKL7	Q9DKL7 spodoptera
699	6	0.7	68	12	O56640	O56640 human cytom	772	6	0.7	95	2	O52341	O52341 escherichia
700	6	0.7	68	12	O56641	O56641 human cytom	773	6	0.7	95	10	Q9SPW9	Q9SPW9 picea glauc
701	6	0.7	68	12	O56643	O56643 human cytom	774	6	0.7	96	10	Q9AXB5	Q9AXB5 oryza sativ
702	6	0.7	69	5	Q9XTX4	Q9XTX4 caenorhabdi	775	6	0.7	97	5	Q9N1R5	Q9N1R5 drosophila
703	6	0.7	69	11	Q99JY9	Q99JY9 mus musculus	776	6	0.7	97	17	Q9HSU4	Q9HSU4 halobacteri
704	6	0.7	70	12	Q92916	Q92916 human cytom	777	6	0.7	98	8	Q99967	Q99967 hybospis wi
705	6	0.7	70	7	Q9S182	Q9S182 stizostedio	778	6	0.7	98	8	Q9XMU5	Q9XMU5 tetrahymena
706	6	0.7	70	10	Q9LDA6	Q9LDA6 oryza sativ	779	6	0.7	98	8	Q94T01	Q94T01 mugil cepha
707	6	0.7	70	12	Q9QPR1	Q9QPR1 hepatitis b	780	6	0.7	99	2	Q68702	Q68702 yersinia pe
708	6	0.7	71	2	P97212	P97212 shewanella	781	6	0.7	99	2	Q932X8	Q932X8 salmonella
709	6	0.7	71	2	O69373	O69373 mariana eub	782	6	0.7	99	6	Q95N01	Q95N01 canis fami
710	6	0.7	71	12	Q9YVY2	Q9YVY2 melanoplus	783	6	0.7	99	8	Q9B976	Q9B976 ceratosen
711	6	0.7	73	2	Q9X567	Q9X567 enterococcu	784	6	0.7	99	12	O91J75	O91J75 hepatitis c
712	6	0.7	73	16	Q9PBX4	Q9PBX4 xylella fas	785	6	0.7	99	16	Q9P9S4	Q9P9S4 xylella fas
713	6	0.7	75	2	Q9AHC9	Q9AHC9 comamonas t	786	6	0.7	99	16	Q9I286	Q9I286 pseudomonas
714	6	0.7	75	2	P71186	P71186 enterobacte	787	6	0.7	100	6	O62850	O62850 oviss aries
715	6	0.7	75	2	Q936Y5	Q936Y5 pseudomonas	788	6	0.7	100	16	O50425	O50425 mycobacteri
716	6	0.7	75	10	Q947H9	Q947H9 verbasum n	789	6	0.7	100	17	O51971	O51971 halobacteri
717	6	0.7	75	10	Q947H8	Q947H8 verbasum t	790	6	0.7	100	17	O9HLP4	O9HLP4 thermoplasm
718	6	0.7	75	10	Q947H7	Q947H7 verbasum n	791	6	0.7	101	4	O14660	O14660 homo sapien
719	6	0.7	75	10	Q947H6	Q947H6 verbasum t	792	6	0.7	101	8	Q9G911	Q9G911 ochromonas
720	6	0.7	75	12	Q99A74	Q99A74 bean pod mo	793	6	0.7	101	16	Q99QY1	Q99QY1 staphylococ
721	6	0.7	76	2	O69369	O69369 mariana eub	794	6	0.7	102	2	Q93BG4	Q93BG4 salmonella
722	6	0.7	76	3	Q9Y868	Q9Y868 aspergillus	795	6	0.7	102	2	Q93BA9	Q93BA9 salmonella
723	6	0.7	76	5	Q9VU93	Q9VU93 drosophila	796	6	0.7	102	9	Q38635	Q38635 bacterioph
724	6	0.7	76	12	O91571	O91571 hepatitis b	797	6	0.7	102	9	Q9T0V4	Q9T0V4 bacterioph
725	6	0.7	77	2	Q9R7U5	Q9R7U5 mariana eub	798	6	0.7	102	9	Q38380	Q38380 bacterioph
726	6	0.7	78	2	O69371	O69371 mariana eub	799	6	0.7	102	17	O28468	O28468 archaeoglob
727	6	0.7	78	2	O69372	O69372 mariana eub	800	6	0.7	103	2	O85916	O85916 sphingomona
728	6	0.7	78	10	Q9SPW6	Q9SPW6 picea maria	801	6	0.7	103	2	Q47833	Q47833 enterococcu
729	6	0.7	79	5	Q25824	Q25824 plasmodium	802	6	0.7	103	2	Q93BB5	Q93BB5 salmonella
730	6	0.7	79	16	Q9PE09	Q9PE09 xylella fas	803	6	0.7	103	10	Q9FLP6	Q9FLP6 arabisopsis
731	6	0.7	80	5	Q26724	Q26724 trypanosoma	804	6	0.7	103	10	Q9LJP7	Q9LJP7 arabisopsis
732	6	0.7	81	6	Q29362	Q29362 sus scrofa	805	6	0.7	103	11	Q9JUF2	Q9JUF2 mus musculu
733	6	0.7	81	12	Q69180	Q69180 human cytom	806	6	0.7	104	10	O04293	O04293 arabisopsis
734	6	0.7	82	12	Q81144	Q81144 hepatitis b	807	6	0.7	104	10	O94825	O94825 solanum tub
735	6	0.7	82	12	Q81146	Q81146 hepatitis b	808	6	0.7	104	10	O93X17	O93X17 solanum tub
736	6	0.7	82	12	Q81148	Q81148 hepatitis b	809	6	0.7	104	13	O9PSX8	O9PSX8 ginglymosto
737	6	0.7	82	12	Q81151	Q81151 hepatitis b	810	6	0.7	105	2	Q9RI19	Q9RI19 yersinia pe
738	6	0.7	82	12	Q81153	Q81153 hepatitis b	811	6	0.7	105	9	Q38426	Q38426 bacterioph
739	6	0.7	82	12	Q81155	Q81155 hepatitis b	812	6	0.7	105	10	Q9SPX1	Q9SPX1 picea abies
740	6	0.7	82	12	Q81156	Q81156 hepatitis b	813	6	0.7	105	10	Q9SPW7	Q9SPW7 picea maria
741	6	0.7	84	16	Q9CD77	Q9CD77 mycobacteri	814	6	0.7	105	16	Q92JM7	Q92JM7 rickettsia
742	6	0.7	85	5	Q17668	Q17668 caenorhabdi	815	6	0.7	105	17	O96Z31	O96Z31 sulfolobus
743	6	0.7	85	10	Q94710	Q94710 verbasum n	816	6	0.7	106	1	P71530	P71530 methanococc
744	6	0.7	85	10	Q93VU2	Q93VU2 antirrhinum	817	6	0.7	106	16	Q9PEH3	Q9PEH3 xylella fas
745	6	0.7	86	10	Q947G6	Q947G6 capsicum an	818	6	0.7	106	16	Q9AC88	Q9AC88 staphylococ
746	6	0.7	86	10	Q947G5	Q947G5 capsicum an	819	6	0.7	106	16	Q9AC88	Q9AC88 staphylococ

820	6	0.7	107	3	Q12023	Q12023 saccharomyc	893	6	0.7	122	17	Q97VT1	Q97vt1 sulfolobus
821	6	0.7	107	16	Q9II195	Q9II195 pseudomonas	894	6	0.7	122	2	Q93BF3	Q93bf3 salmonella
822	6	0.7	108	2	Q93BD2	Q93bd2 salmonella	895	6	0.7	123	2	Q93B93	Q93b93 salmonella
823	6	0.7	108	5	Q97023	Q97023 dugesia jap	896	6	0.7	123	3	Q07355	Q07355 saccharomyc
824	6	0.7	108	8	Q9T247	Q9t247 phytophthor	897	6	0.7	123	4	Q9BRA2	Q9bra2 homo sapien
825	6	0.7	109	2	Q45806	Q45806 clostridium	898	6	0.7	123	9	Q9XJA5	Q9xja5 streptococc
826	6	0.7	109	9	Q9G049	Q9g049 phage phlmh	899	6	0.7	123	9	Q9MCK0	Q9mck0 streptococc
827	6	0.7	109	10	Q42026	Q42026 arabadopsis	900	6	0.7	123	12	Q9EOW7	Q9eow7 hepatitis b
828	6	0.7	110	2	Q9S641	Q9s641 streptococc	901	6	0.7	123	12	Q9EOW5	Q9eow5 hepatitis b
829	6	0.7	110	2	Q93BH2	Q93bh2 salmonella	902	6	0.7	123	12	Q9EOW3	Q9eow3 hepatitis b
830	6	0.7	110	5	Q9GQV9	Q9gqv9 condylactis	903	6	0.7	123	12	Q9EOV3	Q9eov3 hepatitis b
831	6	0.7	110	11	Q9RLW7	Q9rlw7 mus musculu	904	6	0.7	123	12	Q9EOU9	Q9eou9 hepatitis b
832	6	0.7	111	2	Q93BE4	Q93be4 salmonella	905	6	0.7	123	12	Q9EOU7	Q9eou7 hepatitis b
833	6	0.7	111	12	Q81143	Q81143 hepatitis b	906	6	0.7	123	12	Q9EOU3	Q9eou3 hepatitis b
834	6	0.7	111	12	Q81145	Q81145 hepatitis b	907	6	0.7	123	12	Q9EQU1	Q9eou1 hepatitis b
835	6	0.7	111	12	Q81147	Q81147 hepatitis b	908	6	0.7	123	12	Q9EOT9	Q9eot9 hepatitis b
836	6	0.7	111	12	Q81152	Q81152 hepatitis b	909	6	0.7	123	12	Q9EOT7	Q9eot7 hepatitis b
837	6	0.7	111	12	Q81154	Q81154 hepatitis b	910	6	0.7	123	12	Q9EOT3	Q9eot3 hepatitis b
838	6	0.7	111	13	Q90Y71	Q90y71 xenopus lae	911	6	0.7	123	12	Q9E0S9	Q9e0s9 hepatitis b
839	6	0.7	111	16	Q07804	Q07804 mycobacteri	912	6	0.7	123	12	Q9E0S7	Q9e0s7 hepatitis b
840	6	0.7	112	10	Q9C504	Q9c504 arabadopsis	913	6	0.7	123	12	Q9E0R9	Q9e0r9 hepatitis b
841	6	0.7	113	2	Q9WK7	Q9wk7 synechococc	914	6	0.7	123	12	Q9E0R3	Q9e0r3 hepatitis b
842	6	0.7	113	3	Q96UP7	Q96up7 pneumocysti	915	6	0.7	123	12	Q9EQ09	Q9e0q9 hepatitis b
843	6	0.7	113	4	Q96EB4	Q96ee4 homo sapien	916	6	0.7	123	12	Q9EQ07	Q9e0q7 hepatitis b
844	6	0.7	113	5	Q96054	Q96054 samia cynth	917	6	0.7	123	12	Q9EQ03	Q9e0q3 hepatitis b
845	6	0.7	113	11	Q9EQI5	Q9eqi5 mus musculu	918	6	0.7	123	12	Q9EQ01	Q9e0q1 hepatitis b
846	6	0.7	113	12	Q9WIL3	Q9wil3 human cytom	919	6	0.7	123	12	Q9EOP3	Q9eop3 hepatitis b
847	6	0.7	113	12	Q9YKH4	Q9ykh4 human cytom	920	6	0.7	123	12	Q9EOP1	Q9eop1 hepatitis b
848	6	0.7	113	12	Q9YKH3	Q9ykh3 human cytom	921	6	0.7	123	12	Q9EON9	Q9eon9 hepatitis b
849	6	0.7	113	16	Q9YK83	Q9ykt83 vibrio chol	922	6	0.7	123	12	Q9EON5	Q9eon5 hepatitis b
850	6	0.7	113	16	Q9JZ04	Q9jz04 neisseria m	923	6	0.7	123	12	Q9WML2	Q9wml2 avian infec
851	6	0.7	113	16	Q9JU00	Q9ju00 neisseria m	924	6	0.7	123	13	Q91BA9	Q9iba9 oryzias lat
852	6	0.7	114	2	Q93BD4	Q93bd4 salmonella	925	6	0.7	124	2	Q87048	Q87048 vibrio chol
853	6	0.7	114	11	Q9CYW0	Q9cyw0 mus musculu	926	6	0.7	124	2	Q93BG7	Q93bg7 salmonella
854	6	0.7	114	16	Q9KFK2	Q9kfk2 bacillus ha	927	6	0.7	124	2	Q93BC4	Q93bc4 salmonella
855	6	0.7	114	16	Q97K18	Q97k18 clostridium	928	6	0.7	124	2	Q932X1	Q932x1 salmonella
856	6	0.7	115	2	Q93BE2	Q93bf2 salmonella	929	6	0.7	124	3	Q13582	Q13582 saccharomyc
857	6	0.7	115	8	Q21854	Q21854 trachymythe	930	6	0.7	124	5	Q9VCR5	Q9vcr5 drosophila
858	6	0.7	115	8	Q21557	Q21557 oryzomys al	931	6	0.7	124	10	Q9LNA3	Q9lna3 arabadopsis
859	6	0.7	116	2	Q93BE2	Q93be2 salmonella	932	6	0.7	124	10	Q9HPJ8	Q9hjp8 halobacteri
860	6	0.7	116	8	Q94YR7	Q94yr7 engraulis j	933	6	0.7	125	2	Q93BD7	Q93bd7 salmonella
861	6	0.7	116	12	Q9ESA3	Q9esa3 hepatitis b	934	6	0.7	125	2	Q93BC7	Q93bc7 salmonella
862	6	0.7	117	2	P74959	P74959 shewanella	935	6	0.7	125	2	Q93BB8	Q93bb8 salmonella
863	6	0.7	117	2	Q93BH1	Q93bh1 salmonella	936	6	0.7	125	2	Q933R3	Q933r3 salmonella
864	6	0.7	117	2	Q93B84	Q93b84 salmonella	937	6	0.7	125	2	Q933E3	Q933e3 salmonella
865	6	0.7	117	16	Q92GP0	Q92gp0 rickettsia	938	6	0.7	125	5	Q02160	Q02160 caenorhabdi
866	6	0.7	118	2	Q9FDC9	Q9fdg9 streptococc	939	6	0.7	125	11	Q9WY95	Q9wv95 mus musculu
867	6	0.7	118	2	Q93BE7	Q93be7 salmonella	940	6	0.7	125	11	Q9D2E6	Q9d2e6 mus musculu
868	6	0.7	118	2	Q93BD6	Q93bd6 salmonella	941	6	0.7	125	16	P74379	P74379 synechocyst
869	6	0.7	118	2	Q932V6	Q932v6 salmonella	942	6	0.7	125	17	Q9YAY4	Q9yay4 aeropyrum p
870	6	0.7	119	2	Q56737	Q56737 shewanella	943	6	0.7	126	2	Q9R7H8	Q9r7h8 borrelia bu
871	6	0.7	119	17	Q9Y9V6	Q9y9v6 aeropyrum p	944	6	0.7	126	2	Q93B95	Q93b95 salmonella
872	6	0.7	119	17	Q97XR8	Q97xr8 sulfolobus	945	6	0.7	126	2	Q93B86	Q93b86 salmonella
873	6	0.7	120	2	Q9LAL3	Q9lal3 moraxella c	946	6	0.7	126	2	Q93205	Q932q5 salmonella
874	6	0.7	120	2	Q93L55	Q93l55 bacteroides	947	6	0.7	126	2	Q93ZQ0	Q9bzq0 homo sapien
875	6	0.7	120	2	Q93BL1	Q93bl1 salmonella	948	6	0.7	126	5	Q77364	Q77364 plasmodium
876	6	0.7	120	4	Q93BC0	Q93bc0 salmonella	949	6	0.7	126	9	Q9T125	Q9t125 staphylococ
877	6	0.7	120	4	Q96Q82	Q96q82 homo sapien	950	6	0.7	126	10	Q42078	Q42078 arabadopsis
878	6	0.7	120	10	Q9SVF4	Q9svf4 arabadopsis	951	6	0.7	126	16	Q91565	Q91565 pseudomonas
879	6	0.7	120	11	Q9CRL1	Q9crl1 mus musculu	952	6	0.7	126	16	Q99SQ0	Q99sq0 staphylococ
880	6	0.7	120	16	Q9KRA5	Q9kra5 vibrio chol	953	6	0.7	127	2	Q93B99	Q93b99 salmonella
881	6	0.7	121	2	Q933F2	Q933f2 salmonella	954	6	0.7	127	2	Q93B91	Q93b91 salmonella
882	6	0.7	121	5	Q9VKM9	Q9vkm9 drosophila	955	6	0.7	127	2	Q93B87	Q93b87 salmonella
883	6	0.7	121	8	Q96809	Q96809 skeletonema	956	6	0.7	127	2	Q933F7	Q933f7 salmonella
884	6	0.7	121	8	Q9BBM5	Q9bbm5 gentianella	957	6	0.7	127	4	Q9Y5J5	Q9y5j5 homo sapien
885	6	0.7	121	16	Q9KG11	Q9kg11 bacillus ha	958	6	0.7	127	5	Q16907	Q16907 anopheles a
886	6	0.7	122	2	Q93BB6	Q93be6 salmonella	959	6	0.7	127	5	Q16908	Q16908 anopheles a
887	6	0.7	122	2	Q93BD9	Q93bd9 salmonella	960	6	0.7	127	5	Q16909	Q16909 anopheles a
888	6	0.7	122	5	Q9W220	Q9w220 drosophila	961	6	0.7	127	16	Q9CNE1	Q9cne1 pasteurella
889	6	0.7	122	12	Q55730	Q55730 chilo iride-	962	6	0.7	128	2	Q9S637	Q9s637 borrelia bu
890	6	0.7	122	16	Q25081	Q25081 helicobacte	963	6	0.7	128	2	Q93MN0	Q93mn0 mycobacteri
891	6	0.7	122	16	Q9ZMC5	Q9zmc5 helicobacte	964	6	0.7	128	2	Q93BB9	Q93bb9 salmonella
892	6	0.7	122	16	Q9PBS9	Q9pbs9 xylella fas	965	6	0.7	128	2	Q93B85	Q93b85 salmonella

966 6 0.7 128 2 Q933R6 Q933r6 salmonella  
967 6 0.7 128 2 Q933M8 Q933m8 salmonella  
968 6 0.7 128 2 Q933J0 Q933j0 salmonella  
969 6 0.7 128 8 Q954H1 Q954h1 schmidtea 1  
970 6 0.7 128 8 Q954H0 Q954h0 schmidtea 1  
971 6 0.7 128 8 Q94PK2 Q94pk2 schmidtea 1  
972 6 0.7 128 15 Q9QFL4 Q9qfl4 human immun  
973 6 0.7 128 16 P74086 P74086 synechocyst  
974 6 0.7 128 16 Q9RSG8 Q9rsq8 deincococcus  
975 6 0.7 129 2 Q93BG9 Q93bg9 salmonella  
976 6 0.7 129 2 Q93BB7 Q93bb7 salmonella  
977 6 0.7 129 2 Q93BA5 Q93ba5 salmonella  
978 6 0.7 129 2 Q933C7 Q933c7 salmonella  
979 6 0.7 129 5 Q97048 Q97048 pinctada fu  
980 6 0.7 129 5 Q9UAH3 Q9uah3 pinctada fu  
981 6 0.7 129 10 Q9LR14 Q9lrl4 arabidopsis  
982 6 0.7 129 11 Q9CVY3 Q9cvy3 mus musculus  
983 6 0.7 129 16 Q9PC97 Q9pc97 xyiella fas  
984 6 0.7 130 2 Q93BH0 Q93bh0 salmonella  
985 6 0.7 130 2 Q93BG6 Q93bg6 salmonella  
986 6 0.7 130 2 Q93BF0 Q93bf0 salmonella  
987 6 0.7 130 2 Q93BC5 Q93bc5 salmonella  
988 6 0.7 130 4 Q9BYX6 Q9byx6 homo sapien  
989 6 0.7 130 5 Q9S86 Q9s86 caenorhabdi  
990 6 0.7 131 2 Q93BB6 Q93bb6 salmonella  
991 6 0.7 131 2 Q93B97 Q93b97 salmonella  
992 6 0.7 131 2 Q93B94 Q93b94 salmonella  
993 6 0.7 131 2 Q93B83 Q93b83 salmonella  
994 6 0.7 131 2 Q932Y1 Q932y1 salmonella  
995 6 0.7 131 4 Q9UH42 Q9uh42 homo sapien  
996 6 0.7 131 4 Q9S626 Q9s626 homo sapien  
997 6 0.7 131 5 Q9TW98 Q9tw98 pinctada fu  
998 6 0.7 131 5 Q9TVT2 Q9tvt2 pinctada fu  
999 6 0.7 131 5 Q9U260 Q9u260 caenorhabdi  
1000 6 0.7 131 9 Q9B0A4 Q9b0a4 mycobacteri

ALIGNMENTS

RESULT 1  
Q9NYK1 PRELIMINARY; PRT: 1049 AA.  
ID Q9NYK1;  
AC Q9NYK1;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 7.  
GN TLR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RX MEDLINE=20477806; PubMed=11022119;  
RA Du X., Poltorak A., Wei Y., Beutler B.;  
RT "Three novel mammalian toll-like receptors: gene structure,  
RT expression, and evolution."  
RL Eur. Cytokine Netw. 11:362-371(2000).  
DR EMBL: AF240467; AAF60188.1; -.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR\_Cterm.  
DR InterPro: IPR000372; LRR\_Nterm.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR003591; LRR\_typ.  
DR InterPro: IPR000157; TIR.  
DR Pfam: PF00560; LRR; 12.  
DR Pfam: PF01463; LRRCT; 1.  
DR Pfam: PF01582; TIR; 1.  
DR PRINTS: PR000019; LEURICHRPT.  
DR SMART: SM00370; LRR; 4.

DR SMART: SM00082; LRRCT; 1.  
DR SMART: SM00013; LRRNT; 1.  
DR SMART: SM00369; LRR\_TYP; 3.  
DR SMART: SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE. 1049 AA; 120921 MW; 8C701E9E437F2721 CRC64;  
  
Query Match 100.0%; Score 807; DB 4; Length 1049;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 FPKTLPDVTLDVDPKHNHIVDCTDKHLTEIPGGIPTNTNLTNTLTINHIPIIDISPASFHRLD 60  
Db 30 FPKTLPDVTLDVDPKHNHIVDCTDKHLTEIPGGIPTNTNLTNTLTINHIPIIDISPASFHRLD 89  
  
QY 61 HLVEIDFRNCVPIPLGSKNNMICIKRLQIKPRFSGLTYLKSLYLDGNQLLEIPOGLPPS 120  
Db 90 HLVEIDFRNCVPIPLGSKNNMICIKRLQIKPRFSGLTYLKSLYLDGNQLLEIPOGLPPS 149  
  
QY 121 LQLLSLEANNIESIRKENTELANTIELYLGONCYRNPYVYSYTEKDAFNLTKLKV 180  
Db 150 LQLLSLEANNIESIRKENTELANTIELYLGONCYRNPYVYSYTEKDAFNLTKLKV 209  
  
QY 181 SLKDNNTAVPTVLPSTLTETELYNNMTAKIQEDDFNNLNQILDLSCNCPRYNAPEP 240  
Db 210 SLKDNNTAVPTVLPSTLTETELYNNMTAKIQEDDFNNLNQILDLSCNCPRYNAPEP 269  
  
QY 241 CAPCKNSPLQIPVAFADALTELKVLRLHSNSLQHVPPRWFKNINKLQELDLSQNF 300  
Db 270 CAPCKNSPLQIPVAFADALTELKVLRLHSNSLQHVPPRWFKNINKLQELDLSQNF 329  
  
QY 301 IGDAKFLHFLPSLIQDLDSFNELQVYRASMLNSQAFSSLSKLTLIRIGYVFKELKSPN 360  
Db 330 IGDAKFLHFLPSLIQDLDSFNELQVYRASMLNSQAFSSLSKLTLIRIGYVFKELKSPN 389  
  
QY 361 LSPHLNLQNLVLDLGTNFIKIANLSMFKQFKRLKVIDLSVKNISPSPGSDSEVGFCSNAR 420  
Db 390 LSPHLNLQNLVLDLGTNFIKIANLSMFKQFKRLKVIDLSVKNISPSPGSDSEVGFCSNAR 449  
  
QY 421 TSVESYEPQVLEQHYFRDYKARSRCRFKNKEASPMVSNESCYKQGTLDLSKNSIFFVK 480  
Db 450 TSVESYEPQVLEQHYFRDYKARSRCRFKNKEASPMVSNESCYKQGTLDLSKNSIFFVK 509  
  
QY 481 SSDFOHLSFLKCLNLSGNLISOTLANGSEFOPLAEIRYLDFSNRLDLHSTAFEELHKL 540  
Db 510 SSDFOHLSFLKCLNLSGNLISOTLANGSEFOPLAEIRYLDFSNRLDLHSTAFEELHKL 569  
  
QY 541 VLDISSNSHYFQSEGITHTMLNFTKMLKVLQKLMMDNDISSSTRTMESESURTLEFRGN 600  
Db 570 VLDISSNSHYFQSEGITHTMLNFTKMLKVLQKLMMDNDISSSTRTMESESURTLEFRGN 629  
  
QY 601 HLDVLWREGDNRYLQLFKNLLKLEELDTSKNSLSLPSGVFGMPPNKLNSLANKGLKS 660  
Db 630 HLDVLWREGDNRYLQLFKNLLKLEELDTSKNSLSLPSGVFGMPPNKLNSLANKGLKS 689  
  
QY 661 FSWKKLQCLNLETLDLSHNLQTLTPVERLUSNCSRSILKMLILKNNQIRSLTKYFLQDAFOL 720  
Db 690 FSWKKLQCLNLETLDLSHNLQTLTPVERLUSNCSRSILKMLILKNNQIRSLTKYFLQDAFOL 749  
  
QY 721 RYDLSSNNKIOMIKQTSFPENVLNMLKMLLLHHNHLFLCTCDAVWFVWVNHTEVITYLA 780  
Db 750 RYDLSSNNKIOMIKQTSFPENVLNMLKMLLLHHNHLFLCTCDAVWFVWVNHTEVITYLA 809  
  
QY 781 TDVTCVGGCAHKGQSVISLDLYTCELD 807  
Db 810 TDVTCVGGCAHKGQSVISLDLYTCELD 836  
  
RESULT 2  
Q9NR98 PRELIMINARY; PRT: 1049 AA.  
ID Q9NR98  
AC Q9NR98;



DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 7.  
CN TLR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RX MEDLINE=20477807; PubMed=11022120;  
RA Chuang T.H., Ulevitch R.J.;  
RT "Cloning and characterization of a sub-family of human toll-like  
RT receptors: hTLR7, hTLR8 and hTLR9."  
RL Eur. Cytokine Netw. 11:372-378(2000).  
DR EMBL; AF45702; AAF78035.1; -.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_Out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR InterPro; IPR000157; TIR.  
DR Pfam; PF00560; LRR; 12.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; TIR; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 4.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_Typ; 3.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 1049 AA; 120929 MW; 1C77E43B192A86A9 CRC64;

Query Match 86.1%; Score 695; DB 4; Length 1049;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FPKTLPDVTLDVDPKHHVIVDCTDKHLTEIPGGIPTNTNLTINHPIDISPAFHRLD 60  
Db 30 FPKTLPDVTLDVDPKHHVIVDCTDKHLTEIPGGIPTNTNLTINHPIDISPAFHRLD 89  
Qy 61 HLVEIDFRNCVPIPLGSKNNMCIKRLQIKPRFSGLTYLKSLYLDGNQLLEIPQGLPPS 120  
Db 90 HLVEIDFRNCVPIPLGSKNNMCIKRLQIKPRFSGLTYLKSLYLDGNQLLEIPQGLPPS 149  
Qy 121 LQLLSLEANNIFSIRKENLTLANIEILYLGQNCYRRNPVCYVSYSIEKDFAFLNLTCLKVL 180  
Db 150 LQLLSLEANNIFSIRKENLTLANIEILYLGQNCYRRNPVCYVSYSIEKDFAFLNLTCLKVL 209  
Qy 181 SLKNNVTAVTVPSTLTETELYNNTAKIQEDDFNNLNQIQLDLSGNCPCRYNAPFP 240  
Db 210 SLKNNVTAVTVPSTLTETELYNNTAKIQEDDFNNLNQIQLDLSGNCPCRYNAPFP 269  
Qy 241 CAPKNNSPLOIPVNAFDALTEKVLRLHNSLQHPVPRWFKNKQLQELDSONFLAKE 300  
Db 270 CAPKNNSPLOIPVNAFDALTEKVLRLHNSLQHPVPRWFKNKQLQELDSONFLAKE 329  
Qy 301 IGDAKFLHFLPSLIQDLSFNFELQVYRASNMNLSQAFSSLSKLIRGIRGVYFVKELKSFN 360  
Db 330 IGDAKFLHFLPSLIQDLSFNFELQVYRASNMNLSQAFSSLSKLIRGIRGVYFVKELKSFN 389  
Qy 361 LSPHLNQLNLEVDLGTNFIKIANLSMPKQKRLKVIDLSVKNKISPSGDSSEVGFCSNAR 420  
Db 390 LSPHLNQLNLEVDLGTNFIKIANLSMPKQKRLKVIDLSVKNKISPSGDSSEVGFCSNAR 449  
Qy 421 TSVESYEPQVLEQLHYFRYDYKARSCFKNKEASPMNSVNSCYKYGGTLDLSKNSIFFVK 480  
Db 450 TSVESYEPQVLEQLHYFRYDYKARSCFKNKEASPMNSVNSCYKYGGTLDLSKNSIFFVK 509

Qy 481 SSDFQHLFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRDLHLHSTAFELHKL 540  
Db 510 SSDFQHLFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRDLHLHSTAFELHKL 569  
Qy 541 VLDISSNSHYFQSEGITHMLNFTKNLVQLKLMNDNDISSSTRTWSESLRTLEPRGN 600  
Db 570 VLDISSNSHYFQSEGITHMLNFTKNLVQLKLMNDNDISSSTRTWSESLRTLEPRGN 629  
Qy 601 HLDVLRWREGDNRYLQLFKNLKLELDISKNSLSFSGVDFGMPNPKNLSLAKNGLKS 660  
Db 630 HLDVLRWREGDNRYLQLFKNLKLELDISKNSLSFSGVDFGMPNPKNLSLAKNGLKS 689  
Qy 661 FSWKKLQCKLNLETLDLSHNQLTTPPERLSNCSRS 695  
Db 690 FSWKKLQCKLNLETLDLSHNQLTTPPERLSNCSRS 724  
RESULT 3  
ID Q923I1 PRELIMINARY; PRT; 1050 AA.  
AC Q923I1;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 7.  
GN TLR7.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;  
RT "Molecular cloning of murine Toll-Like-Receptor 7."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035889; AAK62676.1; -.  
KW Receptor.  
SQ SEQUENCE 1050 AA; 121836 MW; 495B75DEE849D8EE CRC64;

Query Match 4.2%; Score 34; DB 11; Length 1050;  
Best Local Similarity 100.0%; Pred. No. 3.7e-25;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 774 VTIPYLATDVTGCGAHHKGSVISLDLYTCELD 807  
Db 804 VTIPYLATDVTGCGAHHKGSVISLDLYTCELD 837  
RESULT 4  
ID Q91X17 PRELIMINARY; PRT; 1032 AA.  
AC Q91X17;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 8.  
GN TLR8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=SPLEEN;  
RA Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;  
RT "Molecular cloning of murine Toll-Like Receptor 8."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035890; AAK62677.1; -.  
KW Receptor.  
SQ SEQUENCE 1032 AA; 119338 MW; B2812191D99EE42D CRC64;

Query Match 2.1%; Score 17; DB 11; Length 1032;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 225 LDLSGNCPRCYNAPFC 241  
 DB 246 LDLSGNCPRCYNAPFC 262

RESULT 5  
 Q9H5G9 PRELIMINARY; PRT; 363 AA.  
 AC Q9H5G9;  
 DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE CDNA: FLJ23447 FIS, CLONE HSI03346.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SMALL INTESTINE;  
 RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
 RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
 RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,  
 RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
 RA Isogai T., Sugano S.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK027100; BAB15657.1; -.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR InterPro: IPR003591; LRR\_type.  
 DR Pfam: PF00560; LRR; 8.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00370; LRR; 2.  
 DR SMART: SM00369; LRR\_type; 8.  
 SQ SEQUENCE 363 AA; 39926 MW; DD6999A339228C76 CRC64;

Query Match 1.5%; Score 12; DB 4; Length 363;  
 Best Local Similarity 100.0%; Pred. No. 0.003;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 675 LDLSHNQLTVP 686  
 DB 122 LDLSHNQLTVP 133

RESULT 6  
 Q9SN91 PRELIMINARY; PRT; 1232 AA.  
 AC Q9SN91;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE LEUCINE RICH REPEAT-LIKE PROTEIN.  
 GN FIC12.60 OR AF420140.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terry N., Ardiles W., Buysschaert C., Dasseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Villaroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,  
 RA Mayer K.F.X., Lemcke K., Schueller C.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR [2]

RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-305 FROM N.A.  
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Terry N., Ardiles W., Buysschaert C., Dasseville R., De Clerck R.,  
 RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,  
 RA Gielen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AL022224; CAA18239.1; -.  
 DR EMBL; AL161552; CAB79014.1; -.  
 DR HSSP; P08631; 1AD5.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00560; LRR; 30.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR SMART: SM00370; LRR; 30.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 1232 AA; 135534 MW; EAD6322CE8967726 CRC64;

Query Match 1.5%; Score 12; DB 10; Length 1232;  
 Best Local Similarity 100.0%; Pred. No. 0.0086;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 672 LETLDLSHNQLT 683  
 DB 794 LETLDLSHNQLT 805

RESULT 7  
 Q08817 PRELIMINARY; PRT; 791 AA.  
 AC Q08817;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE CHROMOSOME XV READING FRAME ORF YOR353C.  
 GN YOR353C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Delius H., Hebling U., Hofmann B.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z75261; CAA99682.1; -.  
 DR SGD; S0005880; YOR353C.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR Pfam; PF00560; LRR; 4.

DR SMART; SM00370; LRR; 2.  
SQ SEQUENCE 791 AA; 87325 MW; B0EA559AA4F66199 CRC64;

Query Match 1.2%; Score 10; DB 3; Length 791;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 VLSLKDNNVT 188

Db 141 VLSLKDNNVT 150

RESULT 8

Q9NR97 Q9NR97 PRELIMINARY; PRT; 1041 AA.

AC Q9NR97;

DT 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE TOLL-LIKE RECEPTOR 8.

GN TLR8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=20477807; PubMed=11022120;

RA Chung T.H., Ulevitch R.J.;

RT "Cloning and characterization of a sub-family of human toll-like

receptors: hTLR7, hTLR8 and hTLR9."

RL Eur. Cytokine Netw. 11:372-378(2000).

DR EMBL: AF245703; AAF78036.1; -

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR InterPro; IPR000157; TIR.

DR Pfam; PF00560; LRR; 16

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01582; TIR; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 3.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_TYP; 3.

DR SMART; SM00255; TIR; 1.

KW Receptor.

SQ SEQUENCE 1041 AA; 119827 MW; 39A38B60629291C8 CRC64;

Query Match 1.2%; Score 10; DB 4; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 0.8;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 225 LDLSGNCPRC 234

Db 251 LDLSGNCPRC 260

RESULT 9

Q9NYG9 Q9NYG9 PRELIMINARY; PRT; 1059 AA.

AC Q9NYG9;

DT 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE TOLL-LIKE RECEPTOR 8.

GN TLR8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=20477806; PubMed=11022119;

RA Du X., Poltorak A., Wei Y., Beutler B.;

RT "Three novel mammalian toll-like receptors: gene structure,

expression, and evolution."

RL Eur. Cytokine Netw. 11:362-371(2000).

DR EMBL: AF246971; AAF64061.1; -

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR InterPro; IPR000157; TIR.

DR Pfam; PF00560; LRR; 16

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01582; TIR; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 4.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_TYP; 2.

DR SMART; SM00255; TIR; 1.

KW Receptor.

SQ SEQUENCE 1059 AA; 121751 MW; 2025AEB6DBB7C4AF CRC64;

Query Match 1.2%; Score 10; DB 4; Length 1059;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 225 LDLSGNCPRC 234

Db 269 LDLSGNCPRC 278

RESULT 10

Q63156 Q63156 PRELIMINARY; PRT; 96 AA.

AC Q63156;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE DECORIN (FRAGMENT).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;

RA Moats-Staats B.M., Stiles A.D., Xu L.;

RT "Expression of decorin RNA in rat lung undergoing chronic lung

injury."

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; L75825; AAA85371.1; -

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003592; LRR\_out.

DR Pfam; PF00560; LRR; 3.

DR SMART; SM00370; LRR; 2.

DR NON\_TER 1

FT NON\_TER 96

SQ SEQUENCE 96 AA; 10398 MW; C6D19F1750B050D0 CRC64;

Query Match 1.1%; Score 9; DB 11; Length 96;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IPQGLPPSL 121

Db 12 IPQGLPPSL 20

## RESULT 11

Q9P0Z1 PRELIMINARY; PRT: 250 AA.  
AC Q9P0Z1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE DECORIN B  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CS-Szabo G., Glant T.T.;  
RT "Alternative splicing of human decorin.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF138301; AAF61437.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR Pfam; PF00560; LRR; 3.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00370; LRR; 2.  
DR SMART; SM00013; LRRNT; 1.  
SQ SEQUENCE 250 AA; 27353 MW; 5AA599BE479F68D9 CRC64;

Query Match 1.1%; Score 9; DB 4; Length 250;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121  
|||||  
DB 106 IPQGLPPSL 114

## RESULT 12

Q94L69 PRELIMINARY; PRT: 252 AA.  
AC Q94L69;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE POLYGALACTURONASE INHIBITOR PROTEIN (FRAGMENT).  
GN PGIP.  
OS Potentilla anserina.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Potentilla.  
OX NCBI\_TaxID=57926;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=PO\_AN\_5;  
RT "Phylogenetic relationships among putative genes encoding  
RT polygalacturonase inhibitor proteins (PGIPs) in Rosaceae.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF196916; AAK43430.1; -;  
FT NON\_TER 1  
FT NON\_TER 252  
SQ SEQUENCE 252 AA; 28108 MW; 3CA7578D862DDC6 CRC64;

Query Match 1.1%; Score 9; DB 10; Length 252;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQLT 683  
|||||  
DB 150 LDLSHNQLT 158

## RESULT 13

Q94L68 PRELIMINARY; PRT: 252 AA.  
AC Q94L68;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE POLYGALACTURONASE INHIBITOR PROTEIN (FRAGMENT).  
GN PGIP.  
OS Potentilla anserina.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Potentilla.  
OX NCBI\_TaxID=57926;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=PO\_AN\_7;  
RT Potter D., Oh S.-H., Gao F., Baggett S.;  
RT "Phylogenetic relationships among putative genes encoding  
RT polygalacturonase inhibitor proteins (PGIPs) in Rosaceae.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF196917; AAK43431.1; -;  
FT NON\_TER 1  
FT NON\_TER 252  
SQ SEQUENCE 252 AA; 28078 MW; 39F5C458D80DA380 CRC64;

Query Match 1.1%; Score 9; DB 10; Length 252;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQLT 683  
|||||  
DB 150 LDLSHNQLT 158

## RESULT 14

Q9TTE2 PRELIMINARY; PRT: 360 AA.  
AC Q9TTE2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE DECORIN.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=MYOMETRIUM;  
RX MEDLINE=20113292; PubMed=10644528;  
RA Wu W.X., Zhang Q., Unno N., Derks J.B., Nathanielsz P.W.;  
RT "Characterization of decorin mRNA in pregnant intrauterine tissues of  
RT the ewe and regulation by steroids.";  
RL Am. J. Physiol. 278:C199-C206(2000).  
DR EMBL; AF125041; AAF00585.1; -;  
DR HSSP; P09661; IA9N.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_cyp.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00370; LRR; 2.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TVP; 2.  
SQ SEQUENCE 360 AA; 39972 MW; 0095D0DFDAB8624 CRC64;

Query Match 1.1%; Score 9; DB 6; Length 360;



## RESULT 18

Q921U9 ID Q921U9 PRELIMINARY; PRT; 626 AA.  
AC Q921U9; 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE UNKNOWN (PROTEIN FOR IMAGE:3498778) (FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC010598; AAH10598.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 626 AA; 69123 MW; 5AF3570E270A2DFF CRC64;

Query Match 1.1%; Score 9; DB 11; Length 626;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY 720 LRYLDLSSN 728

|||||||  
Db 246 LRYLDLSSN 254

## RESULT 19

Q9UK78 ID Q9UK78 PRELIMINARY; PRT; 752 AA.  
AC Q9UK78;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE HUMAN TOLL-LIKE RECEPTOR 4 (FRAGMENT).  
GN TLR4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Poltorak A., Smirnova I., Beutler B.;  
RL "Genetic variation at the TLR4 locus.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF177766; AAF07823.1; -.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00560; LRR; 7.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
FT NON\_TER 1  
SQ SEQUENCE 752 AA; 85715 MW; 3275C96C06EA1A2C CRC64;

Query Match 1.1%; Score 9; DB 4; Length 752;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY 723 LDLSSNKIQ 731

|||||||  
Db 93 LDLSSNKIQ 101

## RESULT 20

Q9UM57 ID Q9UM57 PRELIMINARY; PRT; 799 AA.  
AC Q9UM57;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE TOLL-LIKE RECEPTOR 4.  
GN TLR4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-98118556; PubMed=9435236;  
RA Rock F.L., Hardman G., Timans J.C., Kastelein R.A., Bazan J.F.;  
RL "A family of human receptors structurally related to Drosophila Toll.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).  
DR EMBL; U88880; AAC34135.1; -.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR InterPro; IPR000157; TIR.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 2.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 799 AA; 91295 MW; 82F70995E7F2AF9D CRC64;

Query Match 1.1%; Score 9; DB 4; Length 799;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY 723 LDLSSNKIQ 731

|||||||  
Db 140 LDLSSNKIQ 148

## RESULT 21

Q9TSP2 ID Q9TSP2 PRELIMINARY; PRT; 826 AA.  
AC Q9TSP2;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE TOLL-LIKE RECEPTOR 4.  
GN TLR4.  
OS Papio anubis (Olive baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Papio.  
OX NCBI\_TaxID=9555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smirnova I., Poltorak A., Chan E.K.L., Alejos E., McBride C.,  
RA Beutler B.;  
RL "Genetic variation at the TLR4 locus.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF180964; AAF07059.1; -.

DR EMBL; AF180962; AAF07059.1; JOINED.  
DR EMBL; AF180963; AAF07059.1; JOINED.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_typ.  
DR InterPro; IPR000157; TIR.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; TIR; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 1.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 826 AA; 94678 MW; 422777318E5F1769 CRC64;

Query Match 1.1%; Score 9; DB 6; Length 826;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSNKKIQ 731  
DB 180 LDLSNKKIQ 188

RESULT 22  
ID O00206 PRELIMINARY; PRT; 839 AA.  
AC O00206;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE TOLL PROTEIN HOMOLOG (TOLL-LIKE RECEPTOR 4).  
GN TLR4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97379437; PubMed=9237759;  
RA Medzhitov R., Preston-Hurlburt P., Janeway C.A.;  
RT "A human homologue of the Drosophila Toll protein signals activation  
of adaptive immunity."  
RL Nature 388:394-397(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Poltorak A., Smirnova I., Chan E.K.L., Beutler B.;  
RT "Genetic variation at the TLR4 locus."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Arbour N.C., Lorenz E., Schutte B.C., Zabner J., Kline J.N., Jones M.,  
RA Frees K., Watt J.L., Schwartz D.A.;  
RT "A Genetic Basis for a Blunted Response to Endotoxin in Humans."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U93091; AAC80227.1; -.  
DR EMBL; AF17765; AAF05316.1; -.  
DR EMBL; AF172171; AAF89753.1; -.  
DR EMBL; AF172169; AAF89753.1; JOINED.  
DR EMBL; AF172170; AAF89753.1; JOINED.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; TIR; 1.  
DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 2.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 839 AA; 95679 MW; 92C48F55821133E8 CRC64;  
  
Query Match 1.1%; Score 9; DB 4; Length 839;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSNKKIQ 731  
DB 180 LDLSNKKIQ 188

RESULT 23  
ID Q9TTN0 PRELIMINARY; PRT; 839 AA.  
AC Q9TTN0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE TOLL-LIKE RECEPTOR 4.  
GN TLR4.  
OS Pan paniscus (Pygmy chimpanzee) (Homo).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_TaxID=9597;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smirnova I., Poltorak A., Chan E.K.L., Alejos E., McBride C.,  
RA Beutler B.;  
RT "Genetic variation at the TLR4 locus."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF179220; AAF05320.1; -.  
DR EMBL; AF179218; AAF05320.1; JOINED.  
DR EMBL; AF179219; AAF05320.1; JOINED.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_typ.  
DR InterPro; IPR000157; TIR.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; TIR; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 2.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 839 AA; 95637 MW; 3B328C5682127D37 CRC64;

Query Match 1.1%; Score 9; DB 6; Length 839;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSNKKIQ 731  
DB 180 LDLSNKKIQ 188

RESULT 24  
ID Q9LNX8 PRELIMINARY; PRT; 945 AA.  
AC Q9LNX8;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE F2G5.7.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
 RA Kim C., Altafi H., Bei Q., Chin C., Chlou J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RA "Genomic sequence for Arabidopsis thaliana BAC F22C5 from chromosome  
 RT I.";  
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 RL - - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC EMBL; AC022464; AAF79546.1; -;  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR003592; LRR\_out.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR004040; STY\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00560; LRR; 3.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR SMART: SM00370; LRR; 2.  
 DR SMART: SM00221; STYK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 945 AA; 106684 MW; C6DC79C69B69727E CRC64;

Query Match 1.1%; Score 9; DB 10; Length 945;  
 Best Local Similarity 100.0%; Pred. No. 7.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 287 LQELDSLQN 295  
 DB 461 LQELDSLQN 469  
 |||||  
 RESULT 25  
 Q9V701 PRELIMINARY; PRT; 953 AA.  
 ID Q9V701;  
 AC Q9V701;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE CG8561 PROTEIN.  
 GN CG8561.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003815; AAF58265.1; -;  
 DR FLYBase: FBgn0033920; CG8561.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Cterm.  
 DR InterPro: IPR003592; LRR\_out.  
 DR InterPro: IPR003591; LRR\_typ.  
 DR Pfam: PF00560; LRR; 22.  
 DR Pfam: PF01463; LRRCT; 1.  
 DR PRINTS: PR00019; LEURICHRPT.  
 DR SMART: SM00370; LRR; 9.  
 DR SMART: SM00082; LRRCT; 1.  
 DR SMART: SM00369; LRR\_TYP; 3.  
 SQ SEQUENCE 953 AA; 108032 MW; 16D4C22AD854756B CRC64;

Query Match 1.1%; Score 9; DB 5; Length 953;  
 Best Local Similarity 100.0%; Pred. No. 7.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 672 LETLDSLHN 680  
 DB 626 LETLDSLHN 634  
 |||||  
 RESULT 26  
 Q9LKZ4 PRELIMINARY; PRT; 1012 AA.  
 ID Q9LKZ4;  
 AC Q9LKZ4;  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE RECEPTOR-LIKE PROTEIN KINASE 3.  
 GN RLK3.  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21363855; PubMed=11470843;  
 RA Yamamoto E., Knap H.T.;  
 RT "Soybean receptor-like protein kinase genes: paralogous divergence of



```

RT a gene family."
RL Mol. Biol. Evol. 18:1522-1531(2001).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF244890; AAF91324.1; -.
DR HSP; P12931; 1FMK
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR02290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; TYR_pkinase.
DR InterPro; IPR000130; zn_MTPetdse.
DR Pfam; PF00560; LRR; 21.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00370; LRR; 18.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1012 AA; 110322 MW; FD555FB57F99815D CRC64;

Query Match 1.1%; Score 9; DB 10; Length 1012;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 367 LONLEVDL 375
Db 138 LONLEVDL 146
|||||
LONLEVDL 146

RESULT 27
Q9JJ28 PRELIMINARY; PRT; 1271 AA.
AC Q9JJ28;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FLIIH PROTEIN.
GN FLIIH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=20358713; PubMed=10902907;
RA Campbell H.D., Fountain S., Young I.G., Weitz S., Lichter P.,
RA Hoheisel J.D.;
RT "FlIIh, the murine homologue of the Drosophila melanogaster flightless
RT I gene: nucleotide sequence, chromosomal mapping and overlap with
RT LIIH."
RL DNA Seq. 11:29-40(2000).
DR EMBL; AF142329; AAF78453.1; -.
DR HSP; P02640; 2VIL.
DR MGD; MGI:1342286; FlIIh.
DR InterPro; IPR001974; Gelsolin.
DR InterPro; IPR003015; HLH_Myc.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR Pfam; PF00626; Gelsolin; 5.
DR Pfam; PF00560; LRR; 12.
DR PRINTS; PR00597; GELSOLIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00262; GEL; 6.
DR SMART; SM00370; LRR; 7.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 1271 AA; 144802 MW; A9642B10FEF8769 CRC64;

```

```

Query Match 1.1%; Score 9; DB 11; Length 1271;
Best Local Similarity 100.0%; Pred. No. 9; 8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 675 LDLSHNQLT 683
Db 109 LDLSHNQLT 117
|||||
LDLSHNQLT 117

RESULT 28
Q9C6R1 PRELIMINARY; PRT; 1784 AA.
AC Q9C6R1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 201.8 KDA PROTEIN.
GN T18124.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.-L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
DR EMBL; AC079131; AAC50756.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 44.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 20.
KW Hypothetical protein.
SQ SEQUENCE 1784 AA; 201803 MW; 98AEB6FFD6AC8F9D CRC64;

Query Match 1.1%; Score 9; DB 10; Length 1784;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 287 LQELDLSQ 295
Db 226 LQELDLSQ 234
|||||
LQELDLSQ 234

RESULT 29
Q9KJL0 PRELIMINARY; PRT; 226 AA.
AC Q9KJL0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

```

DE HYPOTHETICAL 24.2 KDA PROTEIN.  
OS Trichodesmium sp. IMS101.  
OC Bacteria; Cyanobacteria; Oscillatoriales; Trichodesmium.  
OX NCBI\_TaxID=57878;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dominic B., Zani S., Chen Y.-B., Mellon M.T., Zehr J.P.;  
RT "Organization of the nif genes of the nonheterocystous cyanobacterium  
RL Trichodesmium sp. IMS101."  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF167538; AAF82647.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 226 AA; 24179 MW; 8895C0C95151FE83 CRC64;  
  
Query Match 1.0%; Score 8; DB 2; Length 226;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 254 VNAFDALT 261  
| | | | | | | | | |  
Db 37 VNAFDALT 44  
  
RESULT 30  
Q9NL56 PRELIMINARY; PRT; 336 AA.  
AC  
Q9NL56;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.  
GN GAPDH.  
OS Spirometra erinaceieuropaei.  
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;  
OC Pseudophyllidae; Diphylllobothriidae; Spirometra.  
OX NCBI\_TaxID=99802;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang H., Hirai K., Sato K.;  
RT "Molecular cloning and expression of the gene encoding Spirometra  
RL erinaceieuropaei glyceraldehyde-3-phosphate dehydrogenase."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE  
+ NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.  
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
DEHYDROGENASE FAMILY.  
DR EMBL: AB031067; BAA90773.1; -.  
DR HSSP: P56649; 1SZJ.  
DR InterPro: IPR000173; GAP\_DH.  
DR Pfam: PF00044; gpdh; 1.  
DR Pfam: PF02800; gpdh.C; 1.  
DR PRINTS: PR00078; G3PDHGRGNASE.  
DR PROSITE: PS00071; GAPDH; 1.  
KW Glycolysis; NAD; Oxidoreductase.  
SQ SEQUENCE 336 AA; 35979 MW; 9A61E463B828B44C CRC64;

Query Match 1.0%; Score 8; DB 5; Length 336;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 481 SSDFOHL 488  
| | | | | | | | | |  
Db 283 SSDFOHL 290  
  
RESULT 31  
Q9VED2 PRELIMINARY; PRT; 344 AA.  
ID Q9VED2  
AC Q9VED2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE T24H18\_110.

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE CG14316 PROTEIN.  
GN CG14316.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003720; AAF55493.1; -.  
DR FlyBase: FBgn038567; CG14316.  
SQ SEQUENCE 344 AA; 39318 MW; 3ECA9A947157719C7 CRC64;

Query Match 1.0%; Score 8; DB 5; Length 344;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 490 LKCNLSG 497  
| | | | | | | | | |  
Db 140 LKCNLSG 147  
  
RESULT 32  
Q9LX05 PRELIMINARY; PRT; 371 AA.  
ID Q9LX05  
AC Q9LX05;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 40.0 KDA PROTEIN.  
GN T24H18\_110.

OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Robben J., Grymonprez B., Volckaert G., Bancroft I.,  
RA Meves H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL353013; CAB88258.1; -;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR Pfam; PF00560; LRR; 8.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 6.  
KW Hypothetical protein.  
SQ SEQUENCE 371 AA; 39952 MW; CB216176FB5D1E2A CRC64;

Query Match 1.0%; Score 8; DB 10; Length 371;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 493 LNLGNLI 500  
Db 282 LNLGNLI 289  
|||||

RESULT 33  
Q94BN7 PRELIMINARY; PRT; 371 AA.  
ID Q94BN7  
AC Q94BN7  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 40.0 KDA PROTEIN.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,  
RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,  
RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,  
RA Karlin-Neumann G., Kawai J., Kim C., Koesena E., Lam B., Lin J.,  
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,  
RA Shinohara K., Davis R.W., Ecker J.R., Theologis A.;  
RL "Full Length cDNA of gene T24H18.110/AT5g12940 (GI:7630050).";  
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY039985; AAK64162.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 371 AA; 39968 MW; A7217D6AFB5D1E3F CRC64;

Query Match 1.0%; Score 8; DB 10; Length 371;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 493 LNLGNLI 500  
Db 282 LNLGNLI 289  
|||||

RESULT 34

Q9XBW2 PRELIMINARY; PRT; 428 AA.  
ID Q9XBW2  
AC Q9XBW2  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE IMMUNOREACTIVE 47 KDA ANTIGEN PG97.  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;  
OC Porphyromonas.  
OX NCBI\_TaxID=837;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W50;  
RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,  
RA Hocking D., Webb E.;  
RT "Porphyromonas gingivalis polypeptides and nucleic acids.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF153770; AAD38982.1; -;  
SQ SEQUENCE 428 AA; 47149 MW; 18B0F2CA35B7DD13 CRC64;

Query Match 1.0%; Score 8; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 468 TLDLSKNS 475  
Db 236 TLDLSKNS 243  
|||||

RESULT 35  
Q9CXD9 PRELIMINARY; PRT; 443 AA.  
ID Q9CXD9  
AC Q9CXD9  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE 6130400C22RIK PROTEIN.  
GN 6130400C22RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=THYMUS;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK018071; BAB31060.1; -;  
DR MGD; MGI:1921761; 6130400C22RIK.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR\_Cterm.  
DR InterPro: IPR003591; LRR\_typ.

```

DR Pfam; PF00560; LRR; 5.
DR Pfam; PF01463; LRRCT; 2.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00082; LRRCT; 2.
DR SMART; SM00369; LRR_TY; 5.
SQ SEQUENCE 443 AA; 51851 MW; CF5C962262BB555E CRC64;

Query Match
Best Local Similarity 1.0%; Score 8; DB 11; Length 443;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 723 LDLSNKK 730
Db 275 LDLSNKK 282
|||||

RESULT 36
Q96LI5
ID Q96LI5 PRELIMINARY; PRT; 475 AA.
AC Q96LI5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE CDNA FLJ25459 FIS, CLONE TST09038.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba E., Momiya H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Katsoka R.,
RA Kuga N., Kuroda A., Satoh I., Kanata K., Takami S., Terashima Y.,
RA Watanabe M., Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M.,
RA Kawamura M., Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T.,
RA Sugiyama A., Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK058188; BAB71707.1; -.
SQ SEQUENCE 475 AA; 53791 MW; 725E37DA0749C514 CRC64;

Query Match
Best Local Similarity 1.0%; Score 8; DB 4; Length 475;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 722 YLDLSNKK 729
Db 83 YLDLSNKK 90
|||||

RESULT 37
Q9ZB05
ID Q9ZB05 PRELIMINARY; PRT; 506 AA.
AC Q9ZB05;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PHOSPHORIBOSYLPHOSPHATE AMIDOTRANSFERASE (EC 2.4.2.14).
GN PURF.
OS Lactococcus lactis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99168765; PubMed=10071207;
RX Peltonen T., Mantasala P.;

```

```

RT "Isolation and Characterization of a purC(orf)QLF operon from
RT Lactococcus lactis MG1614.";
RL Mol. Gen. Genet. 261:31-41(1999).
CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
DR EMBL; U64311; AAD12627.1; -.
DR HSSP; P00497; 1GPH.
DR MEROPS; C44.001; -.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR000836; Priboseyltran.
DR InterPro; IPR002375; Pur_Pyr_pr_transf.
DR Pfam; PF00310; GATase_2; 1.
DR Pfam; PF00156; Priboseyltran; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Glycosyltransferase; Transferase
SQ SEQUENCE 506 AA; 55670 MW; 384966055D918590 CRC64;

Query Match
Best Local Similarity 1.0%; Score 8; DB 2; Length 506;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 652 SLAKNGLK 659
Db 499 SLAKNGLK 506
|||||

RESULT 38
Q9EN64
ID Q9EN64 PRELIMINARY; PRT; 564 AA.
AC Q9EN64;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CHITINASE.
OS Spodoptera litura nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZHONGSHAN UNIVERSITY;
RA Hu G., Pang Y., Yang K., Li C.;
RT "Localization, cloning and sequence analysis of the chitinase gene of
RT Spodoptera litura nucleopolyhedrovirus.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RX MEDLINE=21425398; PubMed=11531416;
RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,
RA Yang H.;
RT "Sequence Analysis of the Spodoptera litura Multicapsid
RT Nucleopolyhedrovirus Genome.";
RL Virology 287:391-404(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=G2;
RA Yu J., Wang L., Hu X., Pang Y.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AF246707; AAG02378.1; -.
DR EMBL; AF325155; AAL01727.1; -.
DR HSSP; P07254; ICTN.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 564 AA; 62844 MW; 7EDBD7FF40ABA098 CRC64;

```

Query Match 1.0%; Score 8; DB 12; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 339 SLKSLKIL 346  
 Db 260 SLKSLKIL 267  
 |||||

RESULT 39

O04143 PRELIMINARY; PRT; 581 AA.  
 AC O04143;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE HYPOTHETICAL 63.5 KDA PROTEIN.  
 OS Silene latifolia.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.  
 OX NCBI\_TaxID=37657;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FLOWERBUDS;  
 RA Barbacar N., Hinnisdals S., Farbos I., Moneger F., Lardon A.,  
 RA Delichere C., Mouras A., Negrutiu I.;  
 RT "Isolation of early genes expressed in reproductive organs of the  
 RT dioecious white campion (Silene latifolia) by subtraction cloning  
 using an asexual mutant."  
 RL Plant J. 12:805-817(1997).  
 DR EMBL; Y12529; CAA73132.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003592; LRR\_out.  
 DR InterPro; IPR003591; LRR\_Typ.  
 DR Pfam; PF00560; LRR; 19.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00370; LRR; 5.  
 DR SMART; SM00369; LRR\_Typ; 4.  
 KW Hypothetical protein.  
 SQ SEQUENCE 581 AA; 63516 MW; 52E2D16AD1AA3642 CRC64;

Query Match 1.0%; Score 8; DB 10; Length 581;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 222 LQILDLSG 229  
 Db 454 LQILDLSG 461  
 |||||

RESULT 40

O9FHL8 PRELIMINARY; PRT; 589 AA.  
 AC O9FHL8;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DISEASE RESISTANCE PROTEIN-LIKE.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=95397451; PubMed=10470850;  
 RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
 RA Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.

RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
 RT P1 and TAC clones."  
 RL DNA Res. 6:183-195(1999).  
 DR EMBL; AB018110; BAB09556.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003592; LRR\_out.  
 DR Pfam; PF00560; LRR; 16.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00370; LRR; 9.  
 SQ SEQUENCE 589 AA; 64017 MW; 59E44AE437ECBD7C CRC64;

Query Match 1.0%; Score 8; DB 10; Length 589;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 617 FKNLLKLE 624  
 Db 200 FKNLLKLE 207  
 |||||

RESULT 41

O9ARM5 PRELIMINARY; PRT; 601 AA.  
 AC O9ARM5;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE HYPOTHETICAL 65.5 KDA PROTEIN (FRAGMENT).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21178822; PubMed=11283350;  
 RA Rossberg M., Theres K., Acarkan A., Herrero R., Schmitt T.,  
 RA Schumacher K., Schmitz G., Schmidt R.;  
 RT "Comparative sequence analysis reveals extensive microcolinearity in  
 RT the lateral suppressor regions of the tomato, Arabidopsis and Capsella  
 RT genomes."  
 RL Plant Cell 13:979-988(2001).  
 DR EMBL; AJ303346; CAC36384.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 601 AA; 65513 MW; 4341FF426C67A6DA CRC64;

Query Match 1.0%; Score 8; DB 10; Length 601;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 NLEVLDLG 376  
 Db 383 NLEVLDLG 390  
 |||||

RESULT 42

O9ARF5 PRELIMINARY; PRT; 606 AA.  
 AC O9ARF5;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE HYPOTHETICAL 66.2 KDA PROTEIN.  
 OS Capsella rubella.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Capsella.  
 OX NCBI\_TaxID=81985;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE-21178822; PubMed-11283350;  
RA Rosberg M., Theres K., Acarkan A., Herrero R., Schmitt T.,  
RA Schumacher K., Schmitz G., Schmidt R.;  
RT "Comparative sequence analysis reveals extensive microcolinearity in  
RT the lateral suppressor regions of the tomato, Arabidopsis and Capsella  
RT genomes.";  
RL Plant Cell 13:979-988(2001).  
DR EMBL; AJ303349; CAC36388.1; -.  
DR InterPro; IPR001810; F-box.  
DR Pfam; PF00646; F-box; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 606 AA; 66239 MW; 9E4C22E928806462 CRC64;

Query Match 1.0%; Score 8; DB 10; Length 606;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 NLEVLDLG 376  
|||||  
Db 388 NLEVLDLG 395

## RESULT 43

Q92WC6 PRELIMINARY; PRT; 607 AA.  
AC Q92WC6;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE F20N2.2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucotsid II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,  
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,  
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
RA Ecker J.R.;  
RT "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome  
RT I.";  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Shinn P.P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,  
RA Walker M.M., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,  
RA Gonzalez A.A., Hansen N.N.F., Huizar L.L., Kremenetskaia I.I.,  
RA Lenz C.C., Li J.J., Liu S.S., Luros S.S., Rowley D.D., Schwartz J.J.,  
RA Toriumi M.M., Vysotskaja V.V., Yu G.G., Davis R.R.W.,  
RA Federspiel N.N.A., Theologis A.A., Ecker J.J.R.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,

RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC002328; AAF79512.1; -.  
DR InterPro; IPR001810; F-box.  
DR InterPro; IPR003592; LRR\_out.  
DR Pfam; PF00646; F-box; 2.  
DR SMART; SM00370; LRR; 4.  
SQ SEQUENCE 607 AA; 66259 MW; FC213BC291058FAE CRC64;

Query Match 1.0%; Score 8; DB 10; Length 607;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 369 NLEVLDLG 376  
|||||  
Db 389 NLEVLDLG 396

## RESULT 44

O24437 PRELIMINARY; PRT; 612 AA.  
AC O24437;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE RECEPTOR KINASE-LIKE PROTEIN.  
OS Oryza longistaminata (Long-staminate rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoidae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4528;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=IRBB21;  
RC MEDLINE=96106403; PubMed=8525370;  
RA Song W.Y., Wang G.L., Chen L.L., Kim H.S., Pi L.Y., Holsten T.,  
RA Gardner J., Wang B., Zhai W.X., Zhu L.H., Fauquet C., Ronald P.;  
RT "A receptor kinase-like protein encoded by the rice disease resistance  
RT gene, Xa21.";  
RL Science 270:1804-1806(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN=IRBB21;  
RC MEDLINE=97432142; PubMed=9286106;  
RA Song W.Y., Pi L.Y., Wang G.L., Gardner J., Holsten T., Ronald P.C.;  
RT "Evolution of the rice Xa21 disease resistance gene family.";  
RL Plant Cell 9:1279-1287(1997).  
DR EMBL; U72726; AAB82753.1; -.  
DR InterPro; IPR001611; LRR.  
DR Pfam; PF00560; LRR; 21.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 17.  
RW Kinase.  
SQ SEQUENCE 612 AA; 66380 MW; 9BE54BB84242A91F CRC64;

Query Match 1.0%; Score 8; DB 10; Length 612;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 675 LDLSHNQL 682  
|||||  
Db 156 LDLSHNQL 163

## RESULT 45

Q86486

ID Q86486 PRELIMINARY; PRT; 636 AA.  
AC Q86486;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE F PROTEIN.  
OS Rinderpest virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11241;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EGYPT/84;  
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;  
RT "The complete nucleotide sequence of the fusion protein gene of the  
RT vaccine strain of rinderpest virus: comparison with field virus  
RT isolates";  
RL J. Gen. Virol. 0:0-0(0).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EGYPT/84;  
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;  
RT "Nucleotide sequence comparisons of the fusion protein gene from  
RT virulent and attenuated strains of rinderpest virus";  
RL J. Gen. Virol. 75:3611-3617(1994).  
DR EMBL; Z31655; CAA83481.1; -;  
DR HSP: P04849; LSVF.  
DR InterPro: IPR00776; Fusion\_gly.  
DR Pfam: PF00523; fusion\_gly; 1.  
SQ SEQUENCE 636 AA; 67943 MW; 414E0D990821E378 CRC64;

Query Match 1.0%; Score 8; DB 12; Length 636;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 774 VTIPYLAT 781  
Db 101 VTIPYLAT 108  
|||||||

RESULT 46  
Q9LHN5 PRELIMINARY; PRT; 678 AA.  
AC Q9LHN5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE GB|AAD25817.1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl.  
RT TAC and BAC clones";  
RL DNA Res. 7:217-221(2000).  
DR EMBL; AP002039; BAB03093.1; -;  
DR InterPro: IPR002885; PPR.  
DR Pfam: PF01535; PPR; 7.  
SQ SEQUENCE 678 AA; 76416 MW; 0DF2BE1DB5FE1A9F CRC64;

Query Match 1.0%; Score 8; DB 10; Length 678;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 338 SSLKSLKI 345  
Db 271 SSLKSLKI 278  
|||||||

RESULT 47  
Q93ZW6 PRELIMINARY; PRT; 695 AA.  
AC Q93ZW6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE PUTATIVE DISEASE RESISTANCE PROTEIN (FRAGMENT).  
GN MHK7.15/AT5G40920.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,  
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,  
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,  
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,  
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,  
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,  
RA Theologis A.;  
RT "Full length cDNA of gene MHK7.15/AY5940920 (GI:10177430).";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY056226; AAL07075.1; -;  
FT NON\_TER 695  
SQ SEQUENCE 695 AA; 78517 MW; 212B93FF27975CEB CRC64;

Query Match 1.0%; Score 8; DB 10; Length 695;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 536 LHKLEVL 543  
Db 451 LHKLEVL 458  
|||||||

RESULT 48  
O50027 PRELIMINARY; PRT; 720 AA.  
ID O50027  
AC O50027;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HCR9-4B.  
GN HCR9-4B.  
OS Lycopersicon hirsutum.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=62890;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. CF4;  
RX MEDLINE=98074802; PubMed=9413991;  
RA Parniske M., Hammond-Kosack K.E., Golstein C., Thomas C.M.,  
RA Jones D.A., Harrison K., Wulff B.B., Jones J.D.;  
RT "Novel Disease Resistance Specificities Result From Sequence Exchange  
RT Between Tandemly Repeated Genes At The Cf-4/9 Locus Of Tomato.";

RL Cell 91:821-832(1997).  
 DR EMBL; AJ002235; CAA05266.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003592; LRR\_out.  
 DR Pfam; PF00560; LRR; 20.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00370; LRR; 14.  
 SQ SEQUENCE 720 AA; 80913 MW; 5719A4E8EA4BA332 CRC64;

Query Match 1.0%; Score 8; DB 10; Length 720;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 723 LDLSNKI 730  
 |||||  
 Db 578 LDLSNKI 585

## RESULT 49

Oy 92UH7 PRELIMINARY; PRT; 743 AA.  
 AC Q92UH7;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE PUTATIVE DISEASE RESISTANCE PROTEIN.  
 GN ATG24160.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams C., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana";  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC005967; AAD03377.1; -.  
 DR InterPro; IPR000627; Dioxigenase.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003592; LRR\_out.  
 DR Pfam; PF00560; LRR; 16.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00370; LRR; 12.  
 DR PROSITE; PS00083; INTRADIOL\_DIOXYGENAS; UNKNOWN\_1.  
 SQ SEQUENCE 743 AA; 84786 MW; D32BE707DDBF0EDC CRC64;

Query Match 1.0%; Score 8; DB 10; Length 743;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 675 LDLSHNL 682  
 |||||  
 Db 608 LDLSHNL 615

## RESULT 50

Oy 92S81 PRELIMINARY; PRT; 768 AA.  
 AC Q92S81;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE NLOE.  
 GN HCR9-NLOE.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. MONEYMAKER;  
 RX MEDLINE=99125234; PubMed=9926411;  
 RA Parniske M., Wulff B.B., Bonnema G., Thomas C.M., Jones D.A.,  
 RA Jones J.D.;  
 RT "Homologues of the Cf-9 disease resistance gene (Hcr9s) are present at  
 multiple loci on the short arm of tomato chromosome 1";  
 RL Mol. Plant Microbe Interact. 12:93-102(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. MONEYMAKER;  
 RX MEDLINE=99254130;  
 RA Parniske M., Jones J.D.;  
 RT "Recombination between diverged clusters of the tomato Cf-9 plant  
 disease resistance gene family";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:5850-5855(1999).  
 DR EMBL; AF119040; AAD13303.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003592; LRR\_out.  
 DR Pfam; PF00560; LRR; 17.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00370; LRR; 16.  
 SQ SEQUENCE 768 AA; 85686 MW; EF022C4CD4198D4E CRC64;

Query Match 1.0%; Score 8; DB 10; Length 768;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 221 QLQILDLS 228  
 |||||  
 Db 505 QLQILDLS 512

Search completed: July 17, 2002, 09:50:06  
 Job time: 268 sec







PGS2\_PTIG STANDARD; PRT; 360 AA.  
Q9XSD9; Q9XSH4;  
30-MAY-2000 (Rel. 39, Created)  
30-MAY-2000 (Rel. 39, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
Bone proteoglycan II precursor (PG-S2) (Decorin).  
DCN.  
Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A. (LONG FORM).  
RC STRAIN-YORKSHIRE;  
RA Stephenson S., Schnoke M., Vesely I.;  
RT "Cloning of the porcine decorin gene."  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A. (SHORT FORM).  
RC STRAIN-YORKSHIRE; TISSUE=Aorta;  
RA Stephenson S., Schnoke M., Vesely I.;  
RT "Alternatively spliced version of the porcine decorin gene."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE  
RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-  
BETA (BY SIMILARITY).  
CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER  
CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE  
TISSUE OF ORIGIN.  
CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS  
FAMILY.  
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF125537; AAD23578.1; -;  
CC EMBL; AF140270; AAD33862.1; -;  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR000372; LRR\_Nterm.  
CC InterPro; IPR003592; LRR\_out.  
CC InterPro; IPR003591; LRR\_typ.  
CC Pfam; PF00560; LRR; 9.  
CC Pfam; PF01462; LRRNT; 1.  
CC SMART; SM00370; LRR; 1.  
CC SMART; SM00013; LRRNT; 1.  
CC SMART; SM00369; LRR\_TYP; 2.  
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
Repeat; Leucine-rich repeat; Signal; Alternative splicing.  
FT SIGNAL 1 16  
FT PROPEP 17 30  
FT CHAIN 31 360 BONE PROTEOGLYCAN II.  
FT REPEAT 78 99 LRR 1.  
FT REPEAT 100 123 LRR 2.  
FT REPEAT 124 146 LRR 3.  
FT REPEAT 147 168 LRR 4.  
FT REPEAT 169 194 LRR 5.  
FT REPEAT 195 218 LRR 6.  
FT REPEAT 219 239 LRR 7.  
FT REPEAT 240 263 LRR 8.  
FT REPEAT 264 286 LRR 9.  
FT REPEAT 287 309 LRR 10.  
FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
SIMILARITY).  
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 55 68 BY SIMILARITY.

FT DISULFID 314 347 POTENTIAL.  
FT VARSPLIC 281 318 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 360 AA; 39899 MW; 8573DE8DDEBA7509 CRC64;  
Query Match 1.1%; Score 9; DB 1; Length 360;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 113 IPQGLPPSL 121  
| | | | | | | | | |  
Db 216 IPQGLPPSL 224  
| | | | | | | | | |  
RESULT 9  
PGS2\_RABIT STANDARD; PRT; 360 AA.  
ID PGS2\_RABIT  
AC Q28888; Q28608;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Bone proteoglycan II precursor (PG-S2) (Decorin).  
GN DCN  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cornea;  
RX MEDLINE=9512319; PubMed=7822148;  
RA Zhan Q., Burrows R., Cinton C.;  
RT "Cloning and in situ hybridization of rabbit decorin in corneal  
tissues."  
RL Invest. Ophthalmol. Vis. Sci. 36:206-215(1995).  
RN [2]  
RP SEQUENCE OF 38-358 FROM N.A.  
RC TISSUE=Cartilage;  
RA Hering T.M., Kollar J.;  
RT "The primary structure of rabbit chondrocyte decorin deduced from  
nucleotide sequence."  
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE  
RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-  
BETA.  
CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER  
CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE  
TISSUE OF ORIGIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS  
FAMILY.  
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; S76584; AAB33083.1; -;  
CC EMBL; U03394; AAC04315.1; -;  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR000372; LRR\_Nterm.  
CC InterPro; IPR003592; LRR\_out.  
CC InterPro; IPR003591; LRR\_typ.  
CC Pfam; PF00560; LRR; 9.  
CC Pfam; PF01462; LRRNT; 1.  
CC SMART; SM00370; LRR; 1.  
CC SMART; SM00013; LRRNT; 1.  
CC SMART; SM00369; LRR\_TYP; 1.  
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
Repeat; Leucine-rich repeat; Signal.

CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE  
CC TISSUE OF ORIGIN.  
CC -1- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS  
CC FAMILY.  
CC -1- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by - and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Y00712; CAA68702.1; -  
CC PIR: S06280; S06280.  
CC PIR: B31430; B31430.  
CC InterPro: IPR001611; LRR.  
CC InterPro: IPR000372; LRR\_Nterm.  
CC InterPro: IPR003592; LRR\_out.  
CC InterPro: IPR003591; LRR\_typ.  
CC Pfam: PF00560; LRR; 9.  
CC Pfam: PF01462; LRRNT; 1.  
CC SMART: SM00370; LRR; 2.  
CC SMART: SM00013; LRRNT; 1.  
CC SMART: SM00369; LRR\_TYP; 2.  
CC Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
KW Repeat; Leucine-rich repeat; Signal.  
FT SIGNAL 1 16  
FT PROPEP 17 30  
FT CHAIN 31 360  
FT REPEAT 78 99  
FT REPEAT 100 123  
FT REPEAT 124 146  
FT REPEAT 147 168  
FT REPEAT 169 194  
FT REPEAT 195 218  
FT REPEAT 219 239  
FT REPEAT 240 263  
FT REPEAT 264 286  
FT REPEAT 287 309  
FT CARBOHYD 34 34  
FT O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 212 212  
FT CARBOHYD 263 263  
FT CARBOHYD 304 304  
FT DISULFID 55 68  
FT DISULFID 314 347  
FT BY SIMILARITY.  
SQ SEQUENCE 360 AA; 39837 MW; 71E84DA2D87552C0 CRC64;  
  
Query Match 1.1%; Score 9; DB 1; Length 360;  
Best Local Similarity 100.0%; Pred. No. 0.63; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;  
  
Qy 113 IPOGLPPSL 121  
| | | | | | | | | |  
Db 216 IPOGLPPSL 224  
  
RESULT 7  
PGS2\_CANFA  
ID PGS2\_CANFA STANDARD; PRT; 360 AA.  
AC Q29393;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Bone proteoglycan II precursor (PG-S2) (Decorin).  
GN DCN OR DCNIC  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;

RN SEQUENCE FROM N.A.  
RP Glant T.T.;  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 244-259 FROM N.A.  
RL Venta P.J., Brullette J.A., Yuzbaslyan-Gurkan V., Brewer G.J.;  
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE  
CC RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-  
CC BETA (BY SIMILARITY).  
CC -1- PFM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER  
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE  
CC TISSUE OF ORIGIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS  
CC FAMILY.  
CC -1- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by - and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U83141; AAB51245.1; -  
CC EMBL: L77684; AAA98062.1; -  
CC InterPro: IPR001611; LRR.  
CC InterPro: IPR000372; LRR\_Nterm.  
CC InterPro: IPR003592; LRR\_out.  
CC InterPro: IPR003591; LRR\_typ.  
CC Pfam: PF00560; LRR; 9.  
CC Pfam: PF01462; LRRNT; 1.  
CC SMART: SM00370; LRR; 2.  
CC SMART: SM00013; LRRNT; 1.  
CC SMART: SM00369; LRR\_TYP; 1.  
CC Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
KW Repeat; Leucine-rich repeat; Signal.  
FT SIGNAL 1 16  
FT PROPEP 17 30  
FT CHAIN 31 360  
FT REPEAT 78 99  
FT REPEAT 100 123  
FT REPEAT 124 146  
FT REPEAT 147 168  
FT REPEAT 169 194  
FT REPEAT 195 218  
FT REPEAT 219 239  
FT REPEAT 240 263  
FT REPEAT 264 286  
FT REPEAT 287 309  
FT CARBOHYD 34 34  
FT O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
FT N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 212 212  
FT CARBOHYD 263 263  
FT CARBOHYD 304 304  
FT DISULFID 55 68  
FT DISULFID 314 347  
FT BY SIMILARITY.  
SQ SEQUENCE 360 AA; 39980 MW; 99BEE11A9C812906 CRC64;  
  
Query Match 1.1%; Score 9; DB 1; Length 360;  
Best Local Similarity 100.0%; Pred. No. 0.63; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;  
  
Qy 113 IPOGLPPSL 121  
| | | | | | | | | |  
Db 216 IPOGLPPSL 224  
  
RESULT 8  
PGS2\_PIG

alternatively spliced exons in the 5' untranslated region, and mapping of the gene to chromosome 12q23.";  
 Genomics 15:146-160(1993).  
 [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS A; B; C; D AND E).  
 RA Cs-Szabo G., Giant T.T.;  
 "Alternative splicing of human decorin.";  
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RN SEQUENCE OF 31-50.  
 RX MEDLINE=90073579; PubMed=2590169;  
 RA Roughtley P.J., White R.J.;  
 "Dermatan sulphate proteoglycans of human articular cartilage. The properties of dermatan sulphate proteoglycans I and II.";  
 RT Biochem. J. 262:823-827(1989).  
 [6]  
 RN SEQUENCE OF 31-49.  
 RX MEDLINE=8750839; PubMed=3597437;  
 RA Fisher L.W., Hawkins G.R., Tuross N., Termine J.D.;  
 "Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone.";  
 RT J. Biol. Chem. 262:9702-9708(1987).  
 RL J. Biol. Chem. 262:9702-9708(1987).  
 CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE RATE OF FIBRIL FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-BETA.  
 CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE TISSUE OF ORIGIN.  
 CC -!- ALTERNATIVE PRODUCTS: 5 isoforms; A (shown here), B, C, D and E; are produced by alternative splicing.  
 CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS FAMILY.  
 CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M14219; AAB00774.1; -;  
 DR EMBL; L01131; AAA52301.1; ALT\_SEQ.  
 DR EMBL; L01125; AAA52301.1; JOINED.  
 DR EMBL; L01126; AAA52301.1; JOINED.  
 DR EMBL; L01127; AAA52301.1; JOINED.  
 DR EMBL; L01129; AAA52301.1; JOINED.  
 DR EMBL; L01130; AAA52301.1; JOINED.  
 DR EMBL; M98262; AAB60901.1; -;  
 DR EMBL; AF138300; AAD4713.1; -;  
 DR EMBL; AF138301; AAF61437.1; -;  
 DR EMBL; AF138302; AAD4714.1; -;  
 DR EMBL; AF138303; AAF61438.1; -;  
 DR EMBL; AF138304; AAD4715.1; -;  
 DR PIR; A26476; NBHUC8.  
 DR PIR; S05640; S05640.  
 DR PIR; B28457; B28457.  
 DR PIR; A45016; A45016.  
 DR MIM; I25255; -;  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003592; LRR\_Out.  
 DR InterPro; IPR003591; LRR\_typ.  
 DR Pfam; PF00560; LRR; 9.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR SMART; SM00370; LRR; 3.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 1.  
 KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
 KW Repeat; Leucine-rich repeat; Signal; Alternative splicing;  
 KW Polymorphism.

FT SIGNAL 1 16 POTENTIAL.  
 FT PROPEP 17 30  
 FT CHAIN 31 359 BONE PROTEOGLYCAN II.  
 FT REPEAT 77 98 LRR 1.  
 FT REPEAT 99 122 LRR 2.  
 FT REPEAT 123 145 LRR 3.  
 FT REPEAT 146 167 LRR 4.  
 FT REPEAT 168 193 LRR 5.  
 FT REPEAT 194 217 LRR 6.  
 FT REPEAT 218 238 LRR 7.  
 FT REPEAT 239 262 LRR 8.  
 FT REPEAT 263 285 LRR 9.  
 FT REPEAT 286 308 LRR 10.  
 FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN).  
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 54 67 BY SIMILARITY.  
 FT DISULFID 313 346 MISSING (IN ISOFORM B).  
 FT VARSPLIC 71 179 MISSING (IN ISOFORM C).  
 FT VARSPLIC 73 219 MISSING (IN ISOFORM D).  
 FT VARSPLIC 109 295 LDKV -> CLPS (IN ISOFORM E).  
 FT VARSPLIC 72 75 MISSING (IN ISOFORM E).  
 FT VARSPLIC 76 359 E -> Q (IN DBSNP:1803344).  
 FT VARIANT 273 273 /FTIG-VAR\_011975.  
 FT CONFLICT 37 37 G -> A (IN REF. 6).  
 FT CONFLICT 45 45 D -> P (IN REF. 6).  
 FT SEQUENCE 359 AA; 39746 MW; FF511E871A1A52DD CRC64;  
 SQ  
 Query Match 1.1%; Score 9; DB 1; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 0.63;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 113 IPQGLPPSL 121  
 DB 215 IPQGLPPSL 223  
 RESULT 6  
 PGS2\_BOVIN  
 ID PGS2\_BOVIN STANDARD; PRT; 360 AA.  
 AC P21793;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Bone proteoglycan II precursor (PG-S2) (Decorin).  
 GN DCN.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88133946; PubMed=3435485;  
 RA Day A.A., McQuillan C.I., Termine J.D., Young M.R.;  
 RT "Molecular cloning and sequence analysis of the cDNA for small proteoglycan II of bovine bone.";  
 RL Biochem. J. 248:801-805(1987).  
 RN [2]  
 RP SEQUENCE OF 31-54.  
 RX MEDLINE=89123388; PubMed=2914936;  
 RA Choi H.U., Johnson T.L., Pal S., Tang L.H., Rosenberg L., Neame P.J.;  
 RT "Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII, from bovine articular cartilage and skin isolated by octyl-sepharose chromatography.";  
 RL J. Biol. Chem. 264:2876-2884(1989).  
 CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE RATE OF FIBRIL FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-BETA.  
 CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER

CC MEDLINE=20477807; PubMed=11022120;  
RA "Cloning and characterization of a sub-family of human Toll-like  
RT receptors: hTLR7, hTLR8 and hTLR9";  
RL Eur. Cytokine Netw. 11:372-378(2000).  
CC -!- FUNCTION: Participates in the innate immune response to microbial  
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B  
CC activation, cytokine secretion and the inflammatory response (By  
CC similarity).  
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- TISSUE SPECIFICITY: Detected in brain, heart, lung, liver,  
CC placenta, in monocytes, and at lower levels in Col1c+ immature  
CC dendritic cells.  
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF246971; AAF64061.1; -;  
DR EMBL: AF245703; AAF78036.1; -;  
DR MIM: 300366; -;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR\_Cterm.  
DR InterPro: IPR003592; LRR\_Out.  
DR InterPro: IPR003591; LRR\_Typ.  
DR InterPro: IPR000157; TIR.  
DR Pfam: PF00560; LRR; 16.  
DR Pfam: PF01463; LRRCT; 1.  
DR Pfam: PF01582; TIR; 1.  
DR PRINTS: PR00019; LEURICRPT.  
DR SMART: SM00370; LRR; 3.  
DR SMART: SM00082; LRRCT; 1.  
DR SMART: SM00369; LRR\_Typ; 3.  
DR SMART: SM00255; TIR; 1.  
DR PROSITE: PS0104; TIR; 1.  
DR Receptor: Immune response; Inflammatory response; Signal;  
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.  
FT SIGNAL 1 26  
FT CHAIN 27 1041 TOLL-LIKE RECEPTOR 8.  
FT DOMAIN 27 827 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 828 848 POTENTIAL.  
FT DOMAIN 849 1041 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 61 85 LRR 1.  
FT REPEAT 87 112 LRR 2.  
FT REPEAT 124 147 LRR 3.  
FT REPEAT 169 198 LRR 4.  
FT REPEAT 200 223 LRR 5.  
FT REPEAT 225 244 LRR 6.  
FT REPEAT 246 267 LRR 7.  
FT REPEAT 286 309 LRR 8.  
FT REPEAT 311 334 LRR 9.  
FT REPEAT 336 361 LRR 10.  
FT REPEAT 366 389 LRR 11.  
FT REPEAT 393 416 LRR 12.  
FT REPEAT 418 440 LRR 13.  
FT REPEAT 490 511 LRR 14.  
FT REPEAT 529 552 LRR 15.  
FT REPEAT 554 581 LRR 16.  
FT REPEAT 583 607 LRR 17.  
FT REPEAT 609 630 LRR 18.  
FT REPEAT 638 661 LRR 19.  
FT REPEAT 663 686 LRR 20.  
FT REPEAT 687 710 LRR 21.  
FT REPEAT 711 734 LRR 22.

FT REPEAT 736 758  
FT REPEAT 761 785  
FT DOMAIN 878 1025 TIR.  
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 511 511 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 1 1 M -> MKESSLQSSCSLGGKTKK (IN REF. 1).  
FT CONFLICT 217 217 P -> S (IN REF. 1).  
FT CONFLICT 366 366 L -> P (IN REF. 1).  
FT CONFLICT 867 867 V -> I (IN REF. 1).  
SQ SEQUENCE 1041 AA; 119827 MW; 39A38B60629291C8 CRC64;  
  
Query Match 1.2%; Score 10; DB 1; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 225 LDLSGNCPC 234  
|||||||  
Db 251 LDLSGNCPC 260  
  
RESULT 5  
PGS2\_HUMAN  
ID PGS2\_HUMAN STANDARD; PRT; 359 AA.  
AC P07585; Q9Y5N9; Q9Y5N8; Q9P0Z0; Q9P0Z1;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Bone proteoglycan II precursor (PG-S2) (Decorin) (PG40).  
DCN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87017013; PubMed=3484330;  
RA Krusius T., Ruoslahti E.;  
RT "Primary structure of an extracellular matrix proteoglycan core  
RT protein deduced from cloned cDNA";  
RT Proc. Natl. Acad. Sci. U.S.A. 83:7683-7687(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Lung;  
RX MEDLINE=93162643; PubMed=8432527;  
RA Vetter U., Vogel W., Just W., Young M.F., Fisher L.W.;  
RT "Human decorin gene: Intron-exon junctions and chromosomal  
RT localization";  
RL Genomics 15:161-168(1993).  
RN [3]  
RP SEQUENCE OF 1-70 FROM N.A.  
RX MEDLINE=93162642; PubMed=8432526;  
RA Danielson K.G., Fazio A., Cohen I.R., Cannizzaro L., Iozzo R.V.;  
RT "The human decorin gene: Intron-exon organization, discovery of two

```
FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 800 800 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1050 AA; 121836 MW; 495B75DEE849D8EE CRC64;

Query Match 4.2%; Score 34; DB 1; Length 1050;
Best Local Similarity 100.0%; Pred. No. 8.4e-26;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 774 VTIPYLATDVTGCGAHKGQSVISLDLYTCELD 807
|||||
Db 804 VTIPYLATDVTGCGAHKGQSVISLDLYTCELD 837

RESULT 3
TLR8_MOUSE STANDARD; PRT; 1032 AA.
AC P58682;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Toll-like receptor 8 precursor.
GN TLR8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c; TISSUE=Spleen;
RA Heil F. J., Lipford G. B., Wagner H., Bauer S. M.;
RT "Molecular cloning of murine Toll-Like Receptor 8.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Participates in the innate immune response to microbial
CC agents. Acts via MYD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response (By
CC similarity).
CC -1- SUBUNIT: Binds MYD88 via their respective TIR domains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -1- SIMILARITY: CONTAINS 26 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY035890; AAK62677.1; -.
DR PROSITE; PS50104; TIR; 1.
KW Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 1032 TOLL-LIKE RECEPTOR 8.
FT DOMAIN 24 818 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 819 839 POTENTIAL.
FT DOMAIN 840 1032 CYTOSOLIC (POTENTIAL).
FT REPEAT 41 61
FT REPEAT 62 85 LRR 1.
FT REPEAT 87 109 LRR 2.
FT REPEAT 120 143 LRR 3.
FT REPEAT 145 165 LRR 4.
FT REPEAT 166 194 LRR 5.
FT REPEAT 195 218 LRR 6.
FT REPEAT 220 239 LRR 7.
FT REPEAT 240 267 LRR 8.
FT REPEAT 281 304 LRR 9.
FT REPEAT 306 329 LRR 10.
FT REPEAT 306 329 LRR 11.
```

```
FT REPEAT 331 360 LRR 12.
FT REPEAT 361 384 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 436 LRR 15.
FT REPEAT 471 494 LRR 16.
FT REPEAT 520 543 LRR 17.
FT REPEAT 545 572 LRR 18.
FT REPEAT 574 598 LRR 19.
FT REPEAT 600 621 LRR 20.
FT REPEAT 629 652 LRR 21.
FT REPEAT 654 677 LRR 22.
FT REPEAT 678 701 LRR 23.
FT REPEAT 702 725 LRR 24.
FT REPEAT 727 749 LRR 25.
FT REPEAT 752 776 LRR 26.
FT DOMAIN 869 1016 TIR.
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 581 581 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 743 743 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1032 AA; 119338 MW; B2812191D99EB42D CRC64;

Query Match 2.1%; Score 17; DB 1; Length 1032;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 225 LDLSGNCPCRCYNAPPPC 241
|||||
Db 246 LDLSGNCPCRCYNAPPPC 262

RESULT 4
TLR8_HUMAN STANDARD; PRT; 1041 AA.
AC Q9NR97; Q9NYC9;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Toll-like receptor 8 precursor.
GN TLR8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20477806; PubMed=11022119;
RA Du X., Poltorak A., Wei Y., Beutler B.;
RT "Three novel mammalian Toll-like receptors: gene structure,
RT expression, and evolution.";
RL Eur. Cytokine Netw. 11:362-371(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
```

Query Match		100.0%; Score 807; DB 1; Length 1049;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	FPKTLPCDVTLDVDPKHNHVIDCTDKHLTEIPGGIPTNTNLTINIHIPDISPASFRHLD 60
Db	30	FPKTLPCDVTLDVDPKHNHVIDCTDKHLTEIPGGIPTNTNLTINIHIPDISPASFRHLD 89
Qy	61	HLVEDPRCNCVPIPLGSKNNMCIKRQIKPRFSGLTYLKSLSYLDGNGOLLEIPQGLPPS 120
Db	90	HLVEDPRCNCVPIPLGSKNNMCIKRQIKPRFSGLTYLKSLSYLDGNGOLLEIPQGLPPS 149
Qy	121	LQLLSLEANNIFSIRKENLTLANIEILYLGQNCYRNPCVYSYIEKDAFLNLTKLVL 180
Db	150	LQLLSLEANNIFSIRKENLTLANIEILYLGQNCYRNPCVYSYIEKDAFLNLTKLVL 209
Qy	181	SLKONNVAVTVPILPSTILYLYNNIAKIQEDDFNNLNQQLDILDSGNCPCRYNAPFP 240
Db	210	SLKONNVAVTVPILPSTILYLYNNIAKIQEDDFNNLNQQLDILDSGNCPCRYNAPFP 269
Qy	241	CAPCKNSPLQIPYNAPDALTELKVLRLHNSLQHVPPRFKNINKLQELDLSNFLAKE 300
Db	270	CAPCKNSPLQIPYNAPDALTELKVLRLHNSLQHVPPRFKNINKLQELDLSNFLAKE 329
Qy	301	IGDAKFLHFLPSLIQLOLSFNFELQVYRASMLNSQAFSSLSKLIRGYVFKELKSFN 360
Db	330	IGDAKFLHFLPSLIQLOLSFNFELQVYRASMLNSQAFSSLSKLIRGYVFKELKSFN 389
Qy	361	LSPLHLNQLNLEVLDTGNFIKIANLSMFKQFKRLKYVDLSVKNKISPSGDSSEVGFCSNAR 420
Db	390	LSPLHLNQLNLEVLDTGNFIKIANLSMFKQFKRLKYVDLSVKNKISPSGDSSEVGFCSNAR 449
Qy	421	TSVESYEPQVLEQLHYFRYDYARSCRFKNKEASFMSVNSCYKYGOTLDLSKNSIEFVK 480
Db	450	TSVESYEPQVLEQLHYFRYDYARSCRFKNKEASFMSVNSCYKYGOTLDLSKNSIEFVK 509
Qy	481	SSDFQHLFLKCLNLSGNLSIQTUNGSEFQPLAELRYLDFSNRDLHLHSTAFELHKE 540
Db	510	SSDFQHLFLKCLNLSGNLSIQTUNGSEFQPLAELRYLDFSNRDLHLHSTAFELHKE 569
Qy	541	VLDISSNSHYFQSGITHMLNFTKLNKVLQKLMNDNDISSSTRTWESLSRLEFRGN 600
Db	570	VLDISSNSHYFQSGITHMLNFTKLNKVLQKLMNDNDISSSTRTWESLSRLEFRGN 629
Qy	601	HLDVLRGDNRYLQLFKNLKLEBELDISKNSLSFSGVDFGMPNKLNSLAKNGLKS 660
Db	630	HLDVLRGDNRYLQLFKNLKLEBELDISKNSLSFSGVDFGMPNKLNSLAKNGLKS 689
Qy	661	FSWKKLQCLNLETLDLSHNQLTTVPERLSNCSRSKLNLIKNNQIRSLTKYFLQDAFQL 720
Db	690	FSWKKLQCLNLETLDLSHNQLTTVPERLSNCSRSKLNLIKNNQIRSLTKYFLQDAFQL 749
Qy	721	RYLDSSNKKIQIOMIQTSPENVLNKLKLLHHNRFLCTCDVAVFWVWVHNTEVTIPLYA 780
Db	750	RYLDSSNKKIQIOMIQTSPENVLNKLKLLHHNRFLCTCDVAVFWVWVHNTEVTIPLYA 809
Qy	781	TDVTCVGPAGHKGSVSLDLYTCELD 807
Db	810	TDVTCVGPAGHKGSVSLDLYTCELD 836
RESULT 2		
TLR7_MOUSE		
ID	TLR7_MOUSE	STANDARD; PRT; 1050 AA.
AC	P56881	
DT	01-MAR-2002 (Rel. 41, Created)	
DT	01-MAR-2002 (Rel. 41, Last sequence update)	
DE	01-MAR-2002 (Rel. 41, Last annotation update)	
GN	Toll-like receptor 7 precursor.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID=10090;		
[1]		
SEQUENCE FROM N.A.		
TISSUE=Macrophage;		
Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;		
"Molecular cloning of murine Toll-Like-Receptor 7,"		
Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
-1- FUNCTION: Participates in the Innate Immune response to microbial		
agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B		
activation, cytokine secretion and the inflammatory response (By		
similarity).		
-1- SUBUNIT: Binds MyD88 via their respective TIR domains (By		
similarity).		
-1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).		
-1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.		
-1- SIMILARITY: CONTAINS 1 TIR DOMAIN.		
-1- SIMILARITY: CONTAINS 28 LEUCINE-RICH REPEATS (LRR).		
-----		
This SWISS-PROT entry is copyright. It is produced through a collaboration		
between the Swiss Institute of Bioinformatics and the EMBL outstation -		
the European Bioinformatics Institute. There are no restrictions on its		
use by non-profit institutions as long as its content is in no way		
modified and this statement is not removed. Usage by and for commercial		
entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
-----		
EMBL: AY035889; AAK62676.1; -		
PROSITE; PS50104; TIR; 1.		
KW Receptor; Immune response; Inflammatory response; Signal;		
Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.		
FT SIGNAL	1 26	POTENTIAL.
FT CHAIN	27 1050	TOLL-LIKE RECEPTOR 7.
FT DOMAIN	27 837	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	838 858	POTENTIAL.
FT DOMAIN	859 1050	CYTOPLASMIC (POTENTIAL).
FT REPEAT	42 64	LRR 1.
FT REPEAT	65 87	LRR 2.
FT REPEAT	89 111	LRR 3.
FT REPEAT	126 149	LRR 4.
FT REPEAT	151 170	LRR 5.
FT REPEAT	171 195	LRR 6.
FT REPEAT	203 226	LRR 7.
FT REPEAT	228 247	LRR 8.
FT REPEAT	248 273	LRR 9.
FT REPEAT	275 289	LRR 10.
FT REPEAT	290 312	LRR 11.
FT REPEAT	314 337	LRR 12.
FT REPEAT	339 364	LRR 13.
FT REPEAT	369 392	LRR 14.
FT REPEAT	396 419	LRR 15.
FT REPEAT	421 443	LRR 16.
FT REPEAT	493 516	LRR 17.
FT REPEAT	517 542	LRR 18.
FT REPEAT	543 565	LRR 19.
FT REPEAT	567 589	LRR 20.
FT REPEAT	596 619	LRR 21.
FT REPEAT	620 645	LRR 22.
FT REPEAT	650 673	LRR 23.
FT REPEAT	675 698	LRR 24.
FT REPEAT	699 722	LRR 25.
FT REPEAT	724 746	LRR 26.
FT REPEAT	747 770	LRR 27.
FT REPEAT	773 796	LRR 28.
FT DOMAIN	890 1037	TIR.
FT CARBOHYD	66 66	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	69 69	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	167 167	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	190 190	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	215 215	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	287 287	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	324 324	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	535 535	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	591 591	N-LINKED (GLCNAC. . .) (POTENTIAL).



```
983 6 0.7 730 1 CCT2_HUMAN 060583 homo sapien
984 6 0.7 730 1 G13A_DICDI P34115 dictyosteli
985 6 0.7 732 1 KMBB_DICDI P90648 dictyosteli
986 6 0.7 735 1 CIGB_DICDI Q94481 dictyosteli
987 6 0.7 735 1 YDD7_SCHPO Q10432 schizosacch
988 6 0.7 736 1 CPG2_PORGI P95493 porphyromon
989 6 0.7 738 1 I12R_MOUSE Q08837 mus musculu
990 6 0.7 740 1 GNT5_RAT Q08834 rattus norv
991 6 0.7 741 1 CUL5_CAEEL Q23639 caenorhabdi
992 6 0.7 741 1 GNT5_HUMAN Q09328 homo sapien
993 6 0.7 742 1 NEBL_HUMAN Q9ulj8 homo sapien
994 6 0.7 747 1 YFGF_ECOLI P77172 escherichia
995 6 0.7 748 1 KHL1_HUMAN Q9nr64 homo sapien
996 6 0.7 750 1 PSAA_ARATH P56766 arabidopsis
997 6 0.7 750 1 PSAA_LOTJA P58310 lotus japon
998 6 0.7 750 1 PSAA_MARPO P06406 marchantia
999 6 0.7 750 1 PSAA_MESVI Q9mur8 mesostigma
1000 6 0.7 750 1 PSAA_ORYSA P12155 oryza sativ

ALIGNMENTS

RESULT 1
TLR7_HUMAN
ID TLR7_HUMAN STANDARD; PRT; 1049 AA.
AC Q9NYK1; Q9NR98;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
GN Toll-like receptor 7 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN TISSUE=Placenta;
RP MEDLINE=20477806; PubMed=11022119;
RA Du X., Poltorak A., Wei Y., Beutler B.;
RT "Three novel mammalian Toll-like receptors: gene structure,
RT expression, and evolution.";
RL Eur. Cytokine Netw. 11:362-371(2000).
[2]
RN TISSUE=Placenta;
RP MEDLINE=20477807; PubMed=11022120;
RA Chuang T.-H., Olevitch R.J.;
RT "Cloning and characterization of a sub-family of human Toll-like
RT receptors: hTLR7, hTLR8 and hTLR9.";
RL Eur. Cytokine Netw. 11:372-378(2000).
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response (By
CC similarity).
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: Detected in brain, placenta, spleen, stomach,
CC small intestine, lung and in plasmacytoid pre-dendritic cells.
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -!- SIMILARITY: CONTAINS 27 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
```

```
DR EMBL; AF240467; AAF60188.1; -.
DR EMBL; AF245702; AAF78035.1; -.
DR MIM; 300365; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR000157; TIR.
DR Pfam; PF00560; LRR; 12.
DR Pfam; PF01483; LRRCT; 1.
DR Pfam; PF01582; TIR; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 3.
DR SMART; SM00255; TIR; 1.
DR PROSITE; PS0104; TIR; 1.
DR Receptor; Immune response; Inflammatory response; Signal;
DR Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
KW SIGNAL 1 26
FT CHAIN 27 1049
FT DOMAIN 27 839
FT TRANSMEM 840 860
FT DOMAIN 861 1049
FT REPEAT 43 64
FT REPEAT 65 87
FT REPEAT 110 126
FT REPEAT 127 149
FT REPEAT 151 170
FT REPEAT 171 195
FT REPEAT 203 226
FT REPEAT 228 247
FT REPEAT 248 275
FT REPEAT 289 312
FT REPEAT 314 337
FT REPEAT 339 368
FT REPEAT 369 392
FT REPEAT 396 419
FT REPEAT 421 443
FT REPEAT 492 515
FT REPEAT 516 540
FT REPEAT 541 564
FT REPEAT 566 588
FT REPEAT 595 618
FT REPEAT 619 644
FT REPEAT 649 672
FT REPEAT 674 697
FT REPEAT 698 721
FT REPEAT 723 745
FT REPEAT 746 769
FT REPEAT 772 795
FT DOMAIN 889 1036
FT CARBOHYD 66 66
FT CARBOHYD 69 69
FT CARBOHYD 167 167
FT CARBOHYD 202 202
FT CARBOHYD 215 215
FT CARBOHYD 361 361
FT CARBOHYD 413 413
FT CARBOHYD 488 488
FT CARBOHYD 523 523
FT CARBOHYD 534 534
FT CARBOHYD 590 590
FT CARBOHYD 679 679
FT CARBOHYD 720 720
FT CARBOHYD 799 799
FT CONFLICT 725 725
FT CONFLICT 738 738
FT SEQUENCE 1049 AA; 120921 MW; 8C701E9E437F2721 CRC64;
L -> P (IN REF. 2).
```

837	6	0.7	583	1	SCP3_SCHPO	P41000	schizosacch	910	6	0.7	642	1	FLID_CAMJE	Q9phw6	campylobact
838	6	0.7	585	1	DCE2_HUMAN	Q05329	homo sapien	911	6	0.7	644	1	HS72_LYCES	P27322	lycopersico
839	6	0.7	585	1	DCE2_MOUSE	P48320	mus musculus	912	6	0.7	644	1	VP74_NPVOP	O10365	orgyia pseu
840	6	0.7	585	1	HEX3_ADE02	P33279	human adeno	913	6	0.7	645	1	VP74_NPVAC	P15963	autographa
841	6	0.7	585	1	HEX3_ADE05	P12537	human adeno	914	6	0.7	645	1	VP74_NPVCF	P34053	choristoneu
842	6	0.7	587	1	DHSA_CONBU	P51054	coxiella bu	915	6	0.7	649	1	LONH_METJA	Q58812	methanococc
843	6	0.7	587	1	ENV_SRV1	P04027	simian retr	916	6	0.7	651	1	PRIA_MYCLE	Q9ccq3	mycobacteri
844	6	0.7	587	1	SYT3_MOUSE	Q35681	mus musculus	917	6	0.7	652	1	HS7E_SPIOI	P29357	spinacia ol
845	6	0.7	588	1	ATY1_MOUSE	Q9ctg6	mus musculus	918	6	0.7	653	1	HS72_ARATH	P22954	arabidopsis
846	6	0.7	588	1	BINI_MOUSE	O08539	mus musculus	919	6	0.7	653	1	MTS1_STRSA	P29347	streptococc
847	6	0.7	588	1	BINI_RAT	O08839	rattus norv	920	6	0.7	657	1	HUTH_MOUSE	P35492	mus musculu
848	6	0.7	588	1	SYT3_RAT	P40748	rattus norv	921	6	0.7	657	1	HUTH_RAT	P21213	rattus norv
849	6	0.7	589	1	STE7_CANAL	P46599	candida alb	922	6	0.7	658	1	REP_BUCAP	O51889	buchnera ap
850	6	0.7	589	1	VP40_SCMVC	P16046	simian cyto	923	6	0.7	660	1	ABA2_CAPAN	Q96375	capscium an
851	6	0.7	590	1	BTKL_DROME	P08630	drosophila	924	6	0.7	660	1	ALIA_STRPN	P35592	streptococc
852	6	0.7	591	1	DAK2_SCHPO	O74215	schizosacch	925	6	0.7	661	1	OAT2_RAT	O35913	rattus norv
853	6	0.7	591	1	RIB2_YEAST	Q12362	saccharomyc	926	6	0.7	662	1	DCHS_HUMAN	P19113	homo sapien
854	6	0.7	591	1	VGLC_VZVS	P10241	varicella-z	927	6	0.7	662	1	GARP_HUMAN	Q14392	homo sapien
855	6	0.7	591	1	YN48_YEAST	P42846	saccharomyc	928	6	0.7	664	1	DHSA_HUMAN	P31040	homo sapien
856	6	0.7	593	1	BINI_HUMAN	O00499	homo sapien	929	6	0.7	665	1	DHSA_BOVIN	P31039	bos taurus
857	6	0.7	593	1	NDC1_RABIT	Q28615	oryctolagus	930	6	0.7	669	1	OAT3_RAT	P70502	rattus norv
858	6	0.7	595	1	MUTL_RICPR	Q9zc88	rickettsia	931	6	0.7	670	1	OAT3_RAT	O88397	rattus norv
859	6	0.7	596	1	DHSA_RICCN	Q92197	rickettsia	932	6	0.7	671	1	COAT_FCVF9	P27406	feline call
860	6	0.7	596	1	DHSA_RICPR	P31038	rickettsia	933	6	0.7	672	1	GYS_CAEEL	Q9u209	caenorhabdi
861	6	0.7	596	1	NOLX_RHISN	P55711	rhizobium s	934	6	0.7	672	1	WCF4_SCHPO	P87312	schizosacch
862	6	0.7	597	1	MCM3_ENTHI	Q24849	entamoeba h	935	6	0.7	674	1	RGS9_HUMAN	O75916	homo sapien
863	6	0.7	597	1	PTPX_CANAL	P43078	candida alb	936	6	0.7	674	1	TML1_ARATH	P33543	arabidopsis
864	6	0.7	599	1	MX67_YEAST	Q99257	saccharomyc	937	6	0.7	674	1	VTER_HCMVA	P16732	human cytom
865	6	0.7	599	1	SR68_YEAST	P38687	saccharomyc	938	6	0.7	675	1	RGS9_MOUSE	O54828	mus musculu
866	6	0.7	600	1	DHSA_PARDE	Q59661	paracoccus	939	6	0.7	677	1	RPF_C_XANCP	P49246	xanthomonas
867	6	0.7	603	1	ALS_MOUSE	P70389	mus musculus	940	6	0.7	677	1	SKD3_MOUSE	Q60649	mus musculu
868	6	0.7	603	1	ALS_RAT	P35859	rattus norv	941	6	0.7	678	1	GSPD_AERHY	P31780	aeromonas h
869	6	0.7	603	1	US26_HCMVA	P09699	human cytom	942	6	0.7	678	1	GSPD_AERSA	P45778	aeromonas s
870	6	0.7	605	1	ALS_HUMAN	P35858	homo sapien	943	6	0.7	679	1	PAN3_YEAST	P36102	saccharomyc
871	6	0.7	605	1	ALS_PAPHA	O02833	paplo hamad	944	6	0.7	681	1	HS70_PLAFA	P11144	plasmodium
872	6	0.7	605	1	HOP1_YEAST	P20050	saccharomyc	945	6	0.7	681	1	VGP_MABVM	P35253	marburg vir
873	6	0.7	607	1	LCFH_HAEIN	P44446	haemophilus	946	6	0.7	681	1	VGP_MABVP	P35254	marburg vir
874	6	0.7	608	1	ALBU_RAT	P02770	rattus norv	947	6	0.7	681	1	YADD_SCHPO	Q09838	schizosacch
875	6	0.7	608	1	EDD_HELPJ	Q9zkb3	helicobacte	948	6	0.7	682	1	AMPH_CHICK	P50478	gallus gall
876	6	0.7	608	1	EDD_HELPY	P56111	helicobacte	949	6	0.7	682	1	PILJ_PSEAE	P42257	pseudomonas
877	6	0.7	615	1	MUTL_ECOLI	P23367	escherichia	950	6	0.7	683	1	AMPH_RAT	O08838	rattus norv
878	6	0.7	617	1	CYG2_HUMAN	O75343	homo sapien	951	6	0.7	686	1	HS70_PLACB	Q05746	plasmodium
879	6	0.7	618	1	CHEA_LISIN	Q28242	listeria in	952	6	0.7	687	1	SSY5_YEAST	P47002	saccharomyc
880	6	0.7	618	1	CHEA_LISMO	Q48768	listeria mo	953	6	0.7	687	1	TRFE_ONCKI	P79815	oncorhynch
881	6	0.7	618	1	M3K2_HUMAN	Q9Y2U5	homo sapien	954	6	0.7	691	1	EPG_HELPY	Q9zk24	helicobacte
882	6	0.7	618	1	MUTL_SALTY	P14161	salmonella	955	6	0.7	691	1	EPG_HELPY	P56002	helicobacte
883	6	0.7	619	1	M3K2_MOUSE	Q61083	mus musculus	956	6	0.7	691	1	TOKI_YEAST	P40310	saccharomyc
884	6	0.7	619	1	OM70_NEUCR	P23231	neurospora	957	6	0.7	691	1	YI04_YEAST	P40460	saccharomyc
885	6	0.7	620	1	S6A6_BOVIN	Q9mz34	bos taurus	958	6	0.7	691	1	YI04_YEAST	P49418	homo sapien
886	6	0.7	620	1	S6A6_CANFA	O00589	canis famil	959	6	0.7	695	1	AMPH_HUMAN	Q13773	schizosacch
887	6	0.7	620	1	S6A6_HUMAN	P31641	homo sapien	960	6	0.7	699	1	YQQA_CAEEL	Q09299	caenorhabdi
888	6	0.7	621	1	S6A6_MOUSE	O35316	mus musculus	961	6	0.7	702	1	ADA3_YEAST	P32494	saccharomyc
889	6	0.7	621	1	S6A6_MUSCO	P31642	mus cookii	962	6	0.7	702	1	HEL5_AERPE	Q9Yfq8	aeropyrum p
890	6	0.7	621	1	S6A6_RAT	P31643	rattus norv	963	6	0.7	703	1	YKT6_CAEEL	P34317	caenorhabdi
891	6	0.7	621	1	Y425_CHLTR	O84432	chlamydia t	964	6	0.7	710	1	IRAI_MOUSE	Q62406	mus musculu
892	6	0.7	622	1	AMT3_CAEEL	Q21565	caenorhabdi	965	6	0.7	711	1	CAD2_LISMO	Q60048	listeria mo
893	6	0.7	622	1	VAIL_TREPA	O83444	treponema p	966	6	0.7	711	1	YCBY_HAEIN	P44524	haemophilus
894	6	0.7	626	1	MAG_MOUSE	P20917	mus musculus	967	6	0.7	712	1	IRAI_HUMAN	P51617	homo sapien
895	6	0.7	627	1	DTNB_HUMAN	O60941	homo sapien	968	6	0.7	715	1	PSAA_HUPSQ	Q9mu11	huperzia sq
896	6	0.7	628	1	DEAD_ECOLI	P23304	escherichia	969	6	0.7	715	1	VGI_SPVIR	P15892	spiroplasma
897	6	0.7	631	1	NTPI_VACCV	P05807	vaccinia vl	970	6	0.7	719	1	AD18_MOUSE	Q9r157	mus musculu
898	6	0.7	631	1	RIB2_HUMAN	P04844	homo sapien	971	6	0.7	719	1	PSAA_ADICA	Q9mu13	adiantum ca
899	6	0.7	631	1	RIB2_RAT	P25235	rattus norv	972	6	0.7	719	1	PSAA_ASPND	Q9mu14	asplenium n
900	6	0.7	632	1	ETFD_SCHPO	P07111	s probabile	973	6	0.7	719	1	PSAA_EQUA	Q9mu16	equisetum p
901	6	0.7	633	1	YR45_CAEEL	O89562	caenorhabdi	974	6	0.7	720	1	DNL1_AQUAE	O66880	aquifex aeo
902	6	0.7	637	1	YHML_YEAST	P38856	saccharomyc	975	6	0.7	720	1	KRE6_YEAST	P32486	saccharomyc
903	6	0.7	638	1	PHOS_HAEIN	P44808	haemophilus	976	6	0.7	720	1	PSAA_MARBO	Q9mu11	marilia bo
904	6	0.7	639	1	PPO_SPIOI	P43310	spinacia ol	977	6	0.7	720	1	PSAA_SEQSE	Q9muk3	sequoia sem
905	6	0.7	640	1	ELL2_HUMAN	O00472	homo sapien	978	6	0.7	720	1	YM52_YEAST	Q04322	saccharomyc
906	6	0.7	640	1	YS44_CAEEL	O09372	caenorhabdi	979	6	0.7	721	1	YI11_YEAST	P40498	saccharomyc
907	6	0.7	641	1	CAN6_HUMAN	Q9Y6q1	homo sapien	980	6	0.7	723	1	YE7A_SCHPO	O14162	schizosacch
908	6	0.7	641	1	CAN6_MOUSE	O35846	mus musculus	981	6	0.7	726	1	RNR_MYCPN	P75529	mycoplasma
909	6	0.7	642	1	DEAD_KLEPN	P33906	klebsiella	982	6	0.7	727	1	NETA_DROME	Q24567	drosophila

691	6	0.7	466	1	DCE_LACLA	Q9c9g20 lactococcus	764	6	0.7	506	1	CATA_DROME	P17336 drosophila
692	6	0.7	467	1	EUTA_ECOLI	P76551 escherichia	765	6	0.7	507	1	DAF_CAVPO	Q60401 cavia porce
693	6	0.7	468	1	GLNA_SALTY	P06201 salmonella	766	6	0.7	509	1	HUTH_PSEPU	P21310 pseudomonas
694	6	0.7	468	1	AMH_SCHPO	P36606 schizosacch	767	6	0.7	511	1	NEK3_MOUSE	Q9r0a5 mus musculu
695	6	0.7	471	1	NAH_TENMO	P56634 tenebrio mo	768	6	0.7	511	1	P60_LISGR	Q01835 listeria gr
696	6	0.7	471	1	NOLX_RHIFR	P33213 rhizobium f	769	6	0.7	511	1	SYS_MOUSE	P26638 mus musculu
697	6	0.7	472	1	ATPB_FERIS	O50341 fervidobact	770	6	0.7	512	1	DFN5_MOUSE	Q9z2d3 mus musculu
698	6	0.7	473	1	LCB1_MOUSE	O35704 mus musculu	771	6	0.7	513	1	YA51_METJA	Q58451 methanococ
699	6	0.7	473	1	MA1R_YEAST	P53338 saccharomyc	772	6	0.7	515	1	ADCA_STRPY	Q9a019 streptococc
700	6	0.7	473	1	VL2_HPV29	P50800 human papil	773	6	0.7	515	1	HMSH_DROME	Q03372 drosophila
701	6	0.7	474	1	DLD2_BACSU	P54533 bacillus su	774	6	0.7	515	1	PPZ_SCHPO	P78968 schizosacch
702	6	0.7	474	1	DLDH_HALYO	Q04829 halobacteri	775	6	0.7	515	1	Y141_HUMAN	Q14134 homo sapien
703	6	0.7	476	1	PURA_WHEAT	O24396 triticum ae	776	6	0.7	517	1	CPN1_RANCA	Q92104 rana catesb
704	6	0.7	477	1	Y481_TREPA	O83494 treponema p	777	6	0.7	518	1	PR31_SCHPO	O42904 schizosacch
705	6	0.7	477	1	Y264_SYNV3	P73436 synechocyst	778	6	0.7	518	1	YA48_MYCPN	P75066 mycoplasma
706	6	0.7	478	1	G6PD_BORBU	O51581 borrelia bu	779	6	0.7	519	1	YC82_ASTLO	P58145 astasia lon
707	6	0.7	478	1	GLYC_MOUSE	P50431 mus musculu	780	6	0.7	520	1	Y120_MYCGE	P47366 mycoplasma
708	6	0.7	478	1	GYRB_CYTAAU	Q91c00 cytophaga a	781	6	0.7	522	1	VNFA_AZOVI	P12627 azotobacter
709	6	0.7	478	1	GYRB_CYTTHU	Q91ck1 cytophaga h	782	6	0.7	525	1	GUAA_MYCTU	Q50729 mycobacteri
710	6	0.7	479	1	EFT2_SOYBN	P46280 glycine max	783	6	0.7	525	1	YBFA_YEAST	P34219 saccharomyc
711	6	0.7	479	1	ICE8_HUMAN	O14790 h caspase-8	784	6	0.7	526	1	CATA_CANFA	O97492 canis famil
712	6	0.7	479	1	Y098_MYCPN	P75535 mycoplasma	785	6	0.7	526	1	CATA_CAVPO	O64405 cavia porce
713	6	0.7	480	1	BIND_ARBPU	P4608 arabacia pun	786	6	0.7	526	1	CATA_MOUSE	P4270 mus musculu
714	6	0.7	482	1	CATA_BORPE	P48062 bordetella	787	6	0.7	526	1	CATA_RAT	P04762 rattus norv
715	6	0.7	482	1	CATA_PSEAE	O52762 pseudomonas	788	6	0.7	526	1	MLO1_ARATH	O49621 arabidopsis
716	6	0.7	482	1	K6B2_HUMAN	Q9ubs0 h ribosomal	789	6	0.7	528	1	C318_DROME	Q9vyg5 drosophila
717	6	0.7	483	1	MRCO_MESAU	Q9wub9 mesocricetu	790	6	0.7	528	1	LY41_AQUAE	O66682 aquifex aeo
718	6	0.7	483	1	TRKH_ECOLI	P21166 escherichia	791	6	0.7	529	1	GUAA_MYCLE	P46810 mycobacteri
719	6	0.7	483	1	TRKH_SALTY	Q91612 salmonella	792	6	0.7	533	1	ANFA_AZOVI	P12626 azotobacter
720	6	0.7	484	1	CATA_PROMI	P42321 proteus mir	793	6	0.7	533	1	MASY_ECOLI	P08997 escherichia
721	6	0.7	484	1	CATF_HUMAN	Q9ubx1 homo sapien	794	6	0.7	533	1	PAL1_YEAST	P40960 saccharomyc
722	6	0.7	484	1	KPYK_CHLPN	Q9z984 chlamydia p	795	6	0.7	533	1	REF2_YEAST	P42073 saccharomyc
723	6	0.7	484	1	PURA_MAIZE	O24578 zea mays (m	796	6	0.7	535	1	C7C3_MAIZE	P93703 zea mays (m
724	6	0.7	484	1	SCRB_VIBAL	P13394 vibrio algi	797	6	0.7	537	1	ACEA_EMENI	P38298 emericella
725	6	0.7	485	1	GATB_BORBU	O51316 borrelia bu	798	6	0.7	537	1	YDU2_SCHPO	O13863 schizosacch
726	6	0.7	485	1	HUNB_CLOAL	Q96785 clogmia alb	799	6	0.7	541	1	ASNH_METJA	Q38516 methanococ
727	6	0.7	486	1	GTR5_RABIT	P46408 oryctolagus	800	6	0.7	543	1	CP5A_CANTR	P10615 candida tro
728	6	0.7	486	1	PRL1_ARATH	O42384 arabidopsis	801	6	0.7	543	1	VP61_NPVAC	Q03209 autographa
729	6	0.7	486	1	UAP1_CANAL	O74933 candida alb	802	6	0.7	543	1	YDE3_SCHPO	Q10437 schizosacch
730	6	0.7	487	1	HR3_DROME	P13396 drosophila	803	6	0.7	546	1	PCCB_SACER	P53003 saccharopol
731	6	0.7	488	1	PD31_ARATH	Q91nu4 arabidopsis	804	6	0.7	547	1	SNGL_YEAST	P46950 saccharomyc
732	6	0.7	489	1	SYK_MYCPN	P75500 mycoplasma	805	6	0.7	547	1	TCPA_TETPY	O15891 tetrahymena
733	6	0.7	489	1	VGLC_HSVTH	P18535 turkey herp	806	6	0.7	547	1	CH60_RICCP	Q9zct7 rickettsia
734	6	0.7	490	1	PURA_ARATH	Q96529 arabidopsis	807	6	0.7	548	1	MAOX_MYCTU	P71880 mycobacteri
735	6	0.7	491	1	STK3_HUMAN	Q13188 homo sapien	808	6	0.7	548	1	PHR1_CANAL	P43076 candida alb
736	6	0.7	492	1	C133_DROME	Q9vgb3 drosophila	809	6	0.7	549	1	GYRA_MYCKA	O49608 mycobacteri
737	6	0.7	492	1	COX1_PHYNE	Q02211 phytophor	810	6	0.7	549	1	Y4KD_RHISN	P55524 rhizobium s
738	6	0.7	492	1	MENE_STAAU	Q53634 staphylococ	811	6	0.7	550	1	CH60_RICCP	Q9zct7 rickettsia
739	6	0.7	493	1	CHIL_RHINI	P29025 rhizopus ni	812	6	0.7	550	1	GYRA_MYCGO	O49467 mycobacteri
740	6	0.7	493	1	Y014_HUMAN	Q15048 homo sapien	813	6	0.7	552	1	DP1B_ECOLI	P77510 escherichia
741	6	0.7	495	1	G6PD_STRPN	O54537 streptococ	814	6	0.7	554	1	GYRA_MYCFV	O49166 mycobacteri
742	6	0.7	495	1	YS02_CAEEL	Q09357 caenorhabdi	815	6	0.7	554	1	YJJK_ECOLI	P37797 escherichia
743	6	0.7	496	1	DPN5_HUMAN	O60443 homo sapien	816	6	0.7	555	1	ILVD_BACHD	Q9K8e4 bacillus ha
744	6	0.7	496	1	KPR5_YEAST	Q12265 saccharomyc	817	6	0.7	556	1	MP1P_EMENI	P30303 emericella
745	6	0.7	496	1	YKAB_CAEEL	P34262 caenorhabdi	818	6	0.7	560	1	GPV_HUMAN	P40197 homo sapien
746	6	0.7	497	1	TRPE_ACICA	P23315 acinetobact	819	6	0.7	560	1	VGLC_VZVD	O40256 varicella-2
747	6	0.7	498	1	AMYB_IPOBA	P10537 ipomoea bat	820	6	0.7	562	1	CHS5_CANAL	O74161 candida alb
748	6	0.7	498	1	C72IL_ARATH	Q9scn2 arabidopsis	821	6	0.7	563	1	SVR_CHLMU	O9pit8 chlamydia m
749	6	0.7	498	1	YJ1L_YEAST	P47026 saccharomyc	822	6	0.7	564	1	HSF2_CHICK	P38530 gallus gall
750	6	0.7	498	1	YK04_MYCTU	Q10852 mycobacteri	823	6	0.7	566	1	GUNG_CLOTHM	Q05332 clostridium
751	6	0.7	499	1	NPXR_HUMAN	O95502 homo sapien	824	6	0.7	567	1	PGTA_HUMAN	Q92696 homo sapien
752	6	0.7	500	1	CPN2_MOUSE	P15539 mus musculu	825	6	0.7	567	1	PGTA_RAT	O08602 rattus norv
753	6	0.7	501	1	GLPK_ECOLI	P08859 escherichia	826	6	0.7	570	1	TRM1_YEAST	P15565 s n2,n2-dim
754	6	0.7	501	1	GTR5_HUMAN	P22732 homo sapien	827	6	0.7	570	1	YGR0_YEAST	P53109 saccharomyc
755	6	0.7	501	1	GTR5_MOUSE	Q9wv38 mus musculu	828	6	0.7	571	1	UVRG_MYCNA	O84899 mycoplasma
756	6	0.7	501	1	SUCL_CANAL	P33181 candida alb	829	6	0.7	571	1	UVRG_MYCBV	O84898 mycoplasma
757	6	0.7	502	1	CATA_PIG	O62839 sus scrofa	830	6	0.7	572	1	HLV1_AERSA	O08675 aeromonas s
758	6	0.7	502	1	GTR5_RAT	P43427 rattus norv	831	6	0.7	572	1	HLV2_ILVTV	P23987 infectious
759	6	0.7	503	1	YE29_METJA	Q58824 methanococ	832	6	0.7	574	1	IPA4_SHIFL	P18009 shigella fl
760	6	0.7	505	1	GLPK_PSEAE	Q51390 pseudomonas	833	6	0.7	575	1	RPOC_PLAFA	P14422 plasmodium
761	6	0.7	505	1	STCS_EMENI	Q00714 emericella	834	6	0.7	575	1	TREZ_ARTRM	O9ajn6 arthrobacte
762	6	0.7	506	1	C82_ARATH	Q42602 arabidopsis	835	6	0.7	579	1	Y085_CHLTR	O4087 chlamydia t
763	6	0.7	506	1	CATA_BOVIN	P00432 bos taurus	836	6	0.7	580	1	P69_CHLTR	P15362 mycoplasma

545	1	372	0.7	6	0.7	618	0.7	409	1	GCDH_CAEEL	Q20772	caenorhabdi
546	1	373	0.7	6	0.7	619	0.7	413	1	DEOB_HELPJ	Q9zk37	helicobacte
547	1	373	0.7	6	0.7	620	0.7	414	1	PRSB_NAEFO	Q25544	naegleria f
548	1	373	0.7	6	0.7	621	0.7	415	1	Y012_YEAST	P35193	saccharomyc
549	1	374	0.7	6	0.7	622	0.7	416	1	AIAT_BOVIN	P34955	bos taurus
550	1	374	0.7	6	0.7	623	0.7	416	1	AIAT_SHEEP	P12725	ovis aries
551	1	375	0.7	6	0.7	624	0.7	416	1	AROA_ARCFU	Q28775	archaeoglob
552	1	376	0.7	6	0.7	625	0.7	416	1	O35A_DROME	Q9v3q2	drosophila
553	1	377	0.7	6	0.7	626	0.7	416	1	YEBS_HAEIN	P44387	haemophilus
554	1	377	0.7	6	0.7	627	0.7	417	1	DHMH_PARDE	P29894	paracoccus
555	1	377	0.7	6	0.7	628	0.7	419	1	DCDA_ARCFU	Q29458	archaeoglob
556	1	378	0.7	6	0.7	629	0.7	419	1	FAAA_HUMAN	P16930	homo sapien
557	1	378	0.7	6	0.7	630	0.7	419	1	FAAA_MOUSE	P35505	mus muscuku
558	1	379	0.7	6	0.7	631	0.7	419	1	FAAA_RAT	P25093	rattus norv
559	1	380	0.7	6	0.7	632	0.7	419	1	YHAP_BACSU	Q07523	bacillus su
560	1	380	0.7	6	0.7	633	0.7	421	1	DM3L_MOUSE	Q9cwr8	mus muscuku
561	1	380	0.7	6	0.7	634	0.7	422	1	ADHF_SCHPO	Q09669	schizosacch
562	1	381	0.7	6	0.7	635	0.7	422	1	CSD_BORBU	O51111	borrelia bu
563	1	381	0.7	6	0.7	636	0.7	422	1	GSA_CHLTR	O84212	chlamydia t
564	1	382	0.7	6	0.7	637	0.7	423	1	P65_MYCPN	P53663	mycoplasma
565	1	382	0.7	6	0.7	638	0.7	423	1	YHJ3_YEAST	P38690	saccharomyc
566	1	382	0.7	6	0.7	639	0.7	424	1	GLYA_METTM	P50436	methanobact
567	1	382	0.7	6	0.7	640	0.7	425	1	KCN4_MOUSE	O89109	mus muscuku
568	1	382	0.7	6	0.7	641	0.7	425	1	KCN4_RAT	Q9qyw1	rattus norv
569	1	382	0.7	6	0.7	642	0.7	426	1	BZTB_RHOCA	Q52664	rhodobacter
570	1	383	0.7	6	0.7	643	0.7	427	1	KCN4_HUMAN	O15354	homo sapien
571	1	383	0.7	6	0.7	644	0.7	427	1	PURA_DICDI	P21900	dictyostell
572	1	384	0.7	6	0.7	645	0.7	427	1	Y428_METJA	Q57871	methanococc
573	1	385	0.7	6	0.7	646	0.7	428	1	GLYA_METTH	Q27433	methanobact
574	1	385	0.7	6	0.7	647	0.7	431	1	YM49_YEAST	Q03236	saccharomyc
575	1	385	0.7	6	0.7	648	0.7	431	1	YRY4_CAEEL	Q09354	caenorhabdi
576	1	385	0.7	6	0.7	649	0.7	432	1	DCOR_SCHPO	Q9uq99	schizosacch
577	1	386	0.7	6	0.7	650	0.7	434	1	CENA_WARPO	P12211	marchantia
578	1	386	0.7	6	0.7	651	0.7	434	1	PURA_SCHPO	Q02787	schizosacch
579	1	387	0.7	6	0.7	652	0.7	434	1	YSWJ_CAEEL	O45363	caenorhabdi
580	1	388	0.7	6	0.7	653	0.7	435	1	YTG_C_BACSU	O35024	bacillus su
581	1	389	0.7	6	0.7	654	0.7	436	1	Y177_TREPA	O83207	treponema p
582	1	392	0.7	6	0.7	655	0.7	437	1	YABE_BACSU	P37546	bacillus su
583	1	392	0.7	6	0.7	656	0.7	439	1	MATK_PPIVI	P30071	epifagus vl
584	1	392	0.7	6	0.7	657	0.7	439	1	T2EA_HUMAN	P29083	homo sapien
585	1	392	0.7	6	0.7	658	0.7	441	1	YD40_HAEIN	P44165	haemophilus
586	1	392	0.7	6	0.7	659	0.7	442	1	TUL3_HUMAN	O75386	homo sapien
587	1	393	0.7	6	0.7	660	0.7	445	1	CTB2_HUMAN	P56546	homo sapien
588	1	393	0.7	6	0.7	661	0.7	445	1	CTB2_MOUSE	P56546	mus muscuku
589	1	393	0.7	6	0.7	662	0.7	446	1	SUN_SYNY3	P72943	synecocyst
590	1	393	0.7	6	0.7	663	0.7	446	1	YX71_MYCTU	O50400	mycobacteri
591	1	394	0.7	6	0.7	664	0.7	447	1	AC11_CAEEL	P34429	caenorhabdi
592	1	395	0.7	6	0.7	665	0.7	447	1	MCRA_STRLA	P43485	streptomyce
593	1	395	0.7	6	0.7	666	0.7	448	1	GNTF_BACSU	P12012	bacillus su
594	1	396	0.7	6	0.7	667	0.7	448	1	NCAP_CVBF	P19902	bovine coro
595	1	396	0.7	6	0.7	668	0.7	448	1	NCAP_CVBM	P10527	bovine coro
596	1	396	0.7	6	0.7	669	0.7	448	1	NCAP_CVTKE	P26020	turkey ente
597	1	396	0.7	6	0.7	670	0.7	449	1	A4M1_MOUSE	Q9jfc7	mus muscuku
598	1	396	0.7	6	0.7	671	0.7	449	1	ENGA_MYCPN	P75309	mycoplasma
599	1	396	0.7	6	0.7	672	0.7	449	1	MANB_METJA	Q57842	methanococc
600	1	397	0.7	6	0.7	673	0.7	450	1	Y037_MYCGE	P47283	mycoplasma
601	1	397	0.7	6	0.7	674	0.7	451	1	VAH1_CAEEL	Q20666	caenorhabdi
602	1	397	0.7	6	0.7	675	0.7	451	1	VU10_HSV7J	P52524	human herpe
603	1	398	0.7	6	0.7	676	0.7	453	1	A4M1_HUMAN	O00189	homo sapien
604	1	400	0.7	6	0.7	677	0.7	455	1	YSCP_YERPE	P40395	yersinia pe
605	1	401	0.7	6	0.7	678	0.7	455	1	YXX5_CAEEL	Q18179	caenorhabdi
606	1	402	0.7	6	0.7	679	0.7	456	1	RFK9_ECOLI	P37755	escherichia
607	1	402	0.7	6	0.7	680	0.7	456	1	RINI_PIG	P10775	sus scrofa
608	1	403	0.7	6	0.7	681	0.7	456	1	RINI_RAT	P29315	rattus norv
609	1	403	0.7	6	0.7	682	0.7	457	1	UCR1_YEAST	Q07556	saccharomyc
610	1	404	0.7	6	0.7	683	0.7	459	1	NEK3_HUMAN	P51956	homo sapien
611	1	404	0.7	6	0.7	684	0.7	462	1	LACP_STAXY	Q33814	staphylococ
612	1	405	0.7	6	0.7	685	0.7	465	1	AS3B_HUMAN	Q92485	homo sapien
613	1	405	0.7	6	0.7	686	0.7	465	1	CRU4_BRANA	P33522	brassica na
614	1	406	0.7	6	0.7	687	0.7	465	1	TY3H_SCHMA	O17446	escherisoma
615	1	407	0.7	6	0.7	688	0.7	466	1	DCEA_ECO57	P56228	schistosoma
616	1	408	0.7	6	0.7	689	0.7	466	1	DCEA_ECOLI	P80063	escherichia
617	1	408	0.7	6	0.7	690	0.7	466	1	DCEB_ECOLI	P28302	escherichia

399	6	0.7	311	1	OLF6_RAT	P23267 rattus norv	472	6	0.7	345	1	RDS2_XENLA	O42582 xenopus lae
400	6	0.7	311	1	PLC_STAAR	P45723 staphylococ	473	6	0.7	345	1	RDS_BOVIN	P17810 bos taurus
401	6	0.7	312	1	A2GL_HUMAN	P02750 homo sapien	474	6	0.7	346	1	ARGC_BACSU	P23715 bacillus su
402	6	0.7	313	1	NULM_RHISA	O98824 rhinicephal	475	6	0.7	346	1	RDS_CANFA	P52204 canis famil
403	6	0.7	314	1	YFIF_BACSU	P54722 bacillus su	476	6	0.7	346	1	RDS_FELCA	P35906 felis silve
404	6	0.7	315	1	CPPI_ENTHI	Q01957 entamoeba h	477	6	0.7	346	1	RDS_HUMAN	P23942 homo sapien
405	6	0.7	315	1	YDAA_ECOLI	P03807 escherichia	478	6	0.7	346	1	RDS_MOUSE	P15499 mus musculus
406	6	0.7	316	1	RNH3_UREPA	Q9p993 ureaplasma	479	6	0.7	346	1	RDS_RAT	P17438 rattus norv
407	6	0.7	317	1	ACCO_PEA	P31239 pisum sativ	480	6	0.7	346	1	UL95_HSV7J	P52472 human herpe
408	6	0.7	317	1	APE_PIG	P18650 sus scrofa	481	6	0.7	347	1	NU2M_LAMFL	O21070 lampetra fl
409	6	0.7	317	1	EXOZ_RHIME	P26502 rhizobium m	482	6	0.7	349	1	THI4_ARATH	Q38814 arabidopsis
410	6	0.7	317	1	RPOA_AQAE	O66483 aquifex aeo	483	6	0.7	350	1	LICH_XYLFA	Q9pe46 xylella fas
411	6	0.7	317	1	TLM_MOUSE	P17408 mus musculus	484	6	0.7	350	1	YB04_AQAE	O67189 aquifex aeo
412	6	0.7	317	1	Y402_RICPR	Q9zdc9 rickettsia	485	6	0.7	351	1	DESA_SPIPL	O47994 spirulina p
413	6	0.7	323	1	CYCH_XENLA	P51947 xenopus lae	486	6	0.7	351	1	IHBC_RAT	Q9wuk5 rattus norv
414	6	0.7	323	1	DBDD_HUMAN	Q04828 h trans-1,2	487	6	0.7	352	1	IHBC_MOUSE	P55104 mus musculus
415	6	0.7	323	1	DBDI_HUMAN	P52895 homo sapien	488	6	0.7	352	1	PAP2_CARPA	P14080 carica papa
416	6	0.7	324	1	D3HI_DROME	Q9v8m5 drosophila	489	6	0.7	353	1	FPFS_HUMAN	P44324 homo sapien
417	6	0.7	326	1	MIMI_CHICK	P08940 gallus gall	490	6	0.7	353	1	FPFS_RAT	P05369 r farnesyli
418	6	0.7	326	1	TRXB_BORBU	P94284 borrelia bu	491	6	0.7	354	1	MURG_BUCAI	P57311 bucherna ap
419	6	0.7	326	1	VS09_ROTHT	P10501 human rotav	492	6	0.7	354	1	PGS2_MOUSE	P28654 mus musculus
420	6	0.7	326	1	VS09_ROTHT	P11856 human rotav	493	6	0.7	354	1	PGS2_RAT	Q01129 rattus norv
421	6	0.7	326	1	VS09_ROTHT	P09365 porcine rot	494	6	0.7	354	1	RDS_CHICK	O42281 gallus gall
422	6	0.7	326	1	VS09_ROTHT	P32548 porcine rot	495	6	0.7	355	1	UBPC_HUMAN	O75317 homo sapien
423	6	0.7	326	1	VS09_ROTHT	P27423 porcine rot	496	6	0.7	356	1	COX2_BACP3	Q03438 bacillus ps
424	6	0.7	326	1	Y370_MYCPN	P75230 mycoplasma	497	6	0.7	356	1	MURB_BUCAI	P57153 bucherna ap
425	6	0.7	327	1	GYRA_MYCKE	P27065 mycobacteri	498	6	0.7	357	1	CHEB_BACSU	O05522 bacillus su
426	6	0.7	327	1	YG29_SYNY3	P74346 synechoyst	499	6	0.7	357	1	GDFL_MOUSE	P20863 mus musculus
427	6	0.7	328	1	MDH_AQUAR	Q2zf99 aquaspirill	500	6	0.7	357	1	LDXO_MALDO	P51091 malus domes
428	6	0.7	328	1	Y399_CHLTR	O84404 chlamydia t	501	6	0.7	357	1	PGS2_CHICK	P28075 gallus gall
429	6	0.7	328	1	Y679_CHLMU	Q9pj27 chlamydia m	502	6	0.7	358	1	BI0B_METJA	O58692 methanococc
430	6	0.7	329	1	VS09_ROTCC	P29821 chicken rot	503	6	0.7	358	1	Y701_CHLPN	Q9z7k4 chlamydia p
431	6	0.7	330	1	PGIP_PYRGO	Q05091 pyrus commu	504	6	0.7	360	1	AROC_SALTY	P16280 salmonella
432	6	0.7	331	1	PLIB_AGRKL	O83231 agkistrodon	505	6	0.7	360	1	AROC_SALTY	P58729 salmonella
433	6	0.7	331	1	UL07_YZVD	P09301 varicella-z	506	6	0.7	360	1	MRAY_HAEIN	P45062 haemophilus
434	6	0.7	332	1	FLIM_BACSU	P23453 bacillus su	507	6	0.7	360	1	PGS2_HORSE	O46542 equus cabal
435	6	0.7	332	1	GUNY_ERWCH	P27032 erwinia chr	508	6	0.7	361	1	PLGR_YEAST	P47180 saccharomyc
436	6	0.7	332	1	SD22_SCHPO	P22194 schizosacch	509	6	0.7	362	1	1B19_HUMAN	Q08136 homo sapien
437	6	0.7	334	1	CATL_BOVIN	P25975 bos taurus	510	6	0.7	362	1	AGLK_RHIME	Q9z3r9 rhizobium m
438	6	0.7	334	1	COBT_RHOCA	Q52679 rhodobacter	511	6	0.7	362	1	CNTR_CHICK	P51641 gallus gall
439	6	0.7	334	1	DP0B_HUMAN	P06746 homo sapien	512	6	0.7	364	1	OPSG_CAVPO	Q9r024 gavia porce
440	6	0.7	334	1	HBP2_HUMAN	O75031 homo sapien	513	6	0.7	364	1	RECF_XYLFA	Q9pbel xylella fas
441	6	0.7	335	1	VNAT_PHODV	P35947 phocine dis	514	6	0.7	365	1	OPSR_XENLA	O12948 xenopus lae
442	6	0.7	336	1	RUVB_HELPJ	Q9zm57 helicobacte	515	6	0.7	365	1	RRN9_YEAST	P53437 saccharomyc
443	6	0.7	336	1	RUVB_HELPJ	O25699 helicobacte	516	6	0.7	366	1	CD14_MOUSE	P10810 mus musculus
444	6	0.7	336	1	VG27_BPMU5	Q05234 mycobacteri	517	6	0.7	366	1	PGLB_ASPEL	P41750 aspergillus
445	6	0.7	337	1	HNR_ECOLI	P37055 escherichia	518	6	0.7	366	1	T2AL_ACICA	P24546 acinetobact
446	6	0.7	337	1	SPHX_SYNP7	P39665 synechococc	519	6	0.7	366	1	TRMA_ECOLI	P23003 escherichia
447	6	0.7	338	1	ASG1_ECOLI	P18840 escherichia	520	6	0.7	366	1	TRMA_SALTY	P22038 salmonella
448	6	0.7	338	1	LUM_HUMAN	P51884 homo sapien	521	6	0.7	366	1	YQJQ_CAEEL	Q93169 caenorhabdi
449	6	0.7	338	1	LUM_MOUSE	P51885 mus musculus	522	6	0.7	367	1	INX2_DROME	Q9v427 drosophila
450	6	0.7	338	1	LUM_RAT	P51886 rattus norv	523	6	0.7	367	1	TRMU_NEIMA	Q91tj9 neisseria m
451	6	0.7	338	1	SD22_YEAST	P36047 saccharomyc	524	6	0.7	367	1	TRMU_NEIMB	Q91v16 neisseria m
452	6	0.7	339	1	MBR1_YEAST	P23493 saccharomyc	525	6	0.7	368	1	ISPG_BUCAI	P57374 bucherna ap
453	6	0.7	339	1	TORS_DROME	O77277 drosophila	526	6	0.7	368	1	PGLI_ASPNG	P26213 aspergillus
454	6	0.7	339	1	YMS2_MAIZE	P10579 zea mays (m	527	6	0.7	368	1	PGS1_HUMAN	P21810 homo sapien
455	6	0.7	341	1	ARGC_METJA	Q58496 methanococc	528	6	0.7	368	1	RPSD_STAAM	P26766 staphylococ
456	6	0.7	341	1	K6P2_STRCO	Q91118 streptomyc	529	6	0.7	369	1	FATA_CORSA	O42712 coriandrum
457	6	0.7	341	1	MDHM_ARATH	Q9z06 arabidopsis	530	6	0.7	369	1	OPSR_ANOCA	P41592 anolis caro
458	6	0.7	341	1	MDHM_BRANA	Q45744 brassica na	531	6	0.7	369	1	PGS1_BOVIN	P21809 bos taurus
459	6	0.7	341	1	YF71_PYRHO	O53276 pyrococcus	532	6	0.7	369	1	PGS1_CANFA	O02678 canis famil
460	6	0.7	342	1	AQP7_HUMAN	O14520 homo sapien	533	6	0.7	369	1	PGS1_MOUSE	P28653 mus musculus
461	6	0.7	342	1	LUM_BOVIN	Q05443 bos taurus	534	6	0.7	369	1	PGS1_RAT	P47853 rattus norv
462	6	0.7	342	1	RECL_MYXXA	P48291 myxococcus	535	6	0.7	369	1	PGS1_SHEEP	O46390 ovis aries
463	6	0.7	342	1	XVNA_CALSA	P23556 caldocellum	536	6	0.7	369	1	TIB_PARTE	Q27180 paramesium
464	6	0.7	343	1	LUM_CHICK	P51890 gallus gall	537	6	0.7	369	1	TGT_THEMEA	Q9xlp7 thermotoga
465	6	0.7	344	1	HOLA_HAEIN	P43747 haemophilus	538	6	0.7	370	1	PGLI_PENOL	Q9y834 penicillium
466	6	0.7	344	1	LICH_PSES5	P25276 pseudomonas	539	6	0.7	370	1	Y4A4_HELPY	O25685 helicobacte
467	6	0.7	344	1	PARB_MYCTU	O53595 mycobacteri	540	6	0.7	371	1	CD14_RAT	Q63691 rattus norv
468	6	0.7	344	1	POPA_RALSO	Q9rbs0 ralstonia s	541	6	0.7	371	1	GPR_BACME	P22321 bacillus me
469	6	0.7	345	1	ILVC_BUCUN	Q9aq99 bucherna ap	542	6	0.7	371	1	PROJ_BACSU	O07509 bacillus su
470	6	0.7	345	1	PURA_METJA	Q57981 methanococc	543	6	0.7	371	1	Y147_METJA	Q57611 methanococc
471	6	0.7	345	1	RDS1_XENLA	O42581 xenopus lae	544	6	0.7	372	1	CD14_RABIT	Q28680 oryctolagus

253	1	211	0.7	6	0.7	211	1	CORA_HPBA	P24033	hepatitis b	326	6	0.7	264	1	SPRC_CABEL	P34714	caenorhabdi
254	1	211	0.7	6	0.7	327	1	RECR_AQAE	O67455	aquifex aeo	327	6	0.7	266	1	CEMA_CHLVU	P56349	chlorella v
255	1	212	0.7	6	0.7	328	1	ALKH_HAEIN	P44480	h putative	328	6	0.7	266	1	PLSC_MYCPN	P75479	mycoplasma
256	1	212	0.7	6	0.7	329	1	CORA_HPBVT	O05495	hepatitis b	329	6	0.7	267	1	Y209_METJA	O57662	methanococc
257	1	213	0.7	6	0.7	330	1	RTCO_HUMAN	P17081	homo sapien	330	6	0.7	267	1	TYST_NEIGO	O33380	neisseria g
258	1	214	0.7	6	0.7	331	1	CORA_HPBV9	P17099	hepatitis b	331	6	0.7	267	1	YTXD_BACME	P46826	bacillus me
259	1	214	0.7	6	0.7	332	1	MAAI_CABEL	Q18938	caenorhabdi	332	6	0.7	268	1	AROK_THEAC	O9hle5	thermoplasm
260	1	217	0.7	6	0.7	333	1	GTMS_HUMAN	P46439	homo sapien	333	6	0.7	268	1	YLI2_MYCHO	P43056	mycoplasma
261	1	217	0.7	6	0.7	334	1	HRCR_ERWHE	Q47856	erwinia her	334	6	0.7	269	1	SET_DROME	P53997	drosophila
262	1	220	0.7	6	0.7	335	1	GLT1_ARATH	P92995	arabidopsis	335	6	0.7	270	1	KORE_METJA	P53957	methanococc
263	1	220	0.7	6	0.7	336	1	GLT2_ARATH	Q91mc9	arabidopsis	336	6	0.7	270	1	MOPB_RHOCA	Q08386	rhodobacter
264	1	223	0.7	6	0.7	337	1	Y863_METTH	O26951	methanobact	337	6	0.7	271	1	ATBP_STAAR	P18179	staphylococ
265	1	223	0.7	6	0.7	338	1	YAHN_ECOLI	P75693	escherichia	338	6	0.7	271	1	CRK_DROME	Q9xym0	drosophila
266	1	224	0.7	6	0.7	339	1	PLF1_MOUSE	P04095	mus musculus	339	6	0.7	271	1	THTR_AZOVI	P52197	azotobacter
267	1	224	0.7	6	0.7	340	1	RNH2_PYRAB	O9vla9	pyrococcus	340	6	0.7	271	1	YPCP_YEREN	P31485	yersinia en
268	1	224	0.7	6	0.7	341	1	UNG_BACHD	Q9k682	bacillus ha	341	6	0.7	273	1	CIK2_RABIT	Q09081	oryctolagus
269	1	225	0.7	6	0.7	342	1	PSBI_SCHPO	Q9uqy2	schizosacch	342	6	0.7	274	1	FRT1_MOUSE	P70339	mus musculus
270	1	225	0.7	6	0.7	343	1	Y409_MYCPN	P75187	mycoplasma	343	6	0.7	274	1	US18_HCMVA	P09726	human cytom
271	1	226	0.7	6	0.7	344	1	ATP6_PAPHA	O9zxx9	papio hamad	344	6	0.7	275	1	BLO2_SALTY	P05191	salmonella
272	1	226	0.7	6	0.7	345	1	VHEL_PVXX3	P17780	potato viru	345	6	0.7	275	1	BLOF_PSEAE	O51574	pseudomonas
273	1	227	0.7	6	0.7	346	1	YC17_CAMJE	O9pn78	campylobact	346	6	0.7	276	1	DRN2_BOVIN	P56541	bos taurus
274	1	227	0.7	6	0.7	347	1	IF6_PYRAB	O9uvi6	pyrococcus	347	6	0.7	276	1	TODE_PSEPU	P23133	pseudomonas
275	1	229	0.7	6	0.7	348	1	COX2_ASTPE	Q37411	asterina pe	348	6	0.7	276	1	YGFR_ECOLI	Q46818	escherichia
276	1	229	0.7	6	0.7	349	1	COX2_PARKI	P12701	paracentrot	349	6	0.7	277	1	BPA2_STRAU	P29715	streptomyce
277	1	229	0.7	6	0.7	350	1	COX2_PISOC	P25002	pisaster oc	350	6	0.7	277	1	RSU1_HUMAN	Q15404	homo sapien
278	1	229	0.7	6	0.7	351	1	COX2_STRPU	P15545	strongyloce	351	6	0.7	278	1	PRXC_STRAU	Q31168	streptomyce
279	1	229	0.7	6	0.7	352	1	COX2_SYMST	P29880	symptetrum s	352	6	0.7	279	1	BMRR_BACSU	P39075	bacillus su
280	1	230	0.7	6	0.7	353	1	DAG_ANTMA	Q38732	antirrhinum	353	6	0.7	279	1	FRT1_HUMAN	Q92837	homo sapien
281	1	231	0.7	6	0.7	354	1	YCB2_YEAST	P25556	saccharomyc	354	6	0.7	279	1	YQCD_HAEIN	P44153	haemophilus
282	1	233	0.7	6	0.7	355	1	KITH_MOUSE	P04184	mus musculus	355	6	0.7	280	1	CAGT_HELPJ	Q9zlu5	helicobacte
283	1	233	0.7	6	0.7	356	1	YEG7_YEAST	P39983	saccharomyc	356	6	0.7	280	1	CAGT_HELPY	P97245	helicobacte
284	1	234	0.7	6	0.7	357	1	YX02_CABEL	Q11109	caenorhabdi	357	6	0.7	280	1	MDCB_KLEPN	P71422	klebsiella
285	1	237	0.7	6	0.7	358	1	CFI_CALCH	Q42663	callistephu	358	6	0.7	280	1	Y498_MYCLE	P54581	mycobacteri
286	1	238	0.7	6	0.7	359	1	ARD1_YEAST	P07347	saccharomyc	359	6	0.7	280	1	Y498_MYCTU	Q11163	mycobacteri
287	1	238	0.7	6	0.7	360	1	ATP6_MYTED	Q00224	mytilus edu	360	6	0.7	281	1	XYLF_PSEPU	P23106	pseudomonas
288	1	240	0.7	6	0.7	361	1	RKI_CYACA	Q9tm00	cyamidium c	361	6	0.7	282	1	END4_AQAE	O67551	aquifex aeo
289	1	241	0.7	6	0.7	362	1	KDNM_BPT4	P04531	bacterioph	362	6	0.7	283	1	DMPD_PSESP	P19076	pseudomonas
290	1	243	0.7	6	0.7	363	1	YC23_CYPAP	P31605	cyanophora	363	6	0.7	283	1	ISPE_ECOLI	P24209	escherichia
291	1	244	0.7	6	0.7	364	1	PLFR_MOUSE	P04769	mus musculus	364	6	0.7	283	1	ISPE_SALTY	P30753	salmonella
292	1	245	0.7	6	0.7	365	1	RS2_BACSU	P21454	bacillus su	365	6	0.7	283	1	PDXK_ECOLI	P40191	escherichia
293	1	245	0.7	6	0.7	366	1	TEMI_YEAST	P38987	saccharomyc	366	6	0.7	284	1	GOLI_DROME	Q06003	drosophila
294	1	246	0.7	6	0.7	367	1	MCT1_MOUSE	P11034	mus musculus	367	6	0.7	284	1	SUHA_HUMAN	Q06520	homo sapien
295	1	247	0.7	6	0.7	368	1	LAMP_RAT	P49911	rattus norv	368	6	0.7	284	1	SUHA_MACFA	P52842	macaca fasc
296	1	247	0.7	6	0.7	369	1	MCBF_ECOLI	P05529	escherichia	369	6	0.7	285	1	FOLD_BUCAI	P57557	buchnera ap
297	1	249	0.7	6	0.7	370	1	LEC2_ULEEU	P22973	ulex europe	370	6	0.7	286	1	SGBU_HAEIN	P44990	haemophilus
298	1	249	0.7	6	0.7	371	1	PHAL_HUMAN	P39687	homo sapien	371	6	0.7	288	1	ALF_MYCPN	P75089	mycoplasma
299	1	249	0.7	6	0.7	372	1	YF82_METJA	Q58977	methanococc	372	6	0.7	288	1	CGD2_RAT	Q04827	rattus norv
300	1	250	0.7	6	0.7	373	1	SLB2_XENLA	Q9ygp6	xenopus lae	373	6	0.7	288	1	HIS1_METJA	O58601	methanococc
301	1	252	0.7	6	0.7	374	1	ATP6_MARPO	P26853	marchantia	374	6	0.7	289	1	ATPG_BACFI	P22482	bacillus fi
302	1	252	0.7	6	0.7	375	1	HIS6_RHOCA	Q30724	rhodobacter	375	6	0.7	289	1	CGD2_HUMAN	P30279	homo sapien
303	1	252	0.7	6	0.7	376	1	HIS6_RHOSH	P50937	rhodobacter	376	6	0.7	289	1	CGD2_MOUSE	P30280	mus musculus
304	1	252	0.7	6	0.7	377	1	PIGL_HUMAN	Q9y2b2	homo sapien	377	6	0.7	289	1	SSRL_FUGRU	Q42179	fugu rubrip
305	1	252	0.7	6	0.7	378	1	YAG1_ECOLI	P77300	escherichia	378	6	0.7	290	1	RPOD_AERPE	Q9y053	aeropyrum p
306	1	252	0.7	6	0.7	379	1	YELL_MYCPN	P75373	mycoplasma	379	6	0.7	291	1	ERA_MYCPN	P75210	mycoplasma
307	1	253	0.7	6	0.7	380	1	AROE_THEMEA	Q9wy11	thermotoga	380	6	0.7	292	1	BIAA_PROVU	P52660	proteus vul
308	1	253	0.7	6	0.7	381	1	HMBE_BRARE	Q03356	brachydanio	381	6	0.7	292	1	RBSB_HAEIN	P44737	haemophilus
309	1	253	0.7	6	0.7	382	1	Y007_MYCPN	P75105	mycoplasma	382	6	0.7	293	1	ATP6_MYCPN	O50326	mycoplasma
310	1	253	0.7	6	0.7	383	1	YQXM_BACSU	P40949	bacillus su	383	6	0.7	293	1	URER_PROMI	Q02458	proteus mir
311	1	254	0.7	6	0.7	384	1	CGH2_HSVSA	Q01043	herpesvirus	384	6	0.7	294	1	ARGC_BACST	Q07906	bacillus st
312	1	254	0.7	6	0.7	385	1	RACD_DICDI	P34150	dictyosteli	385	6	0.7	296	1	KHSE_LACLA	Q9c9d7	lactococcus
313	1	255	0.7	6	0.7	386	1	TPIS_BUCAI	P57393	buchnera ap	386	6	0.7	297	1	VRPR_SALDU	P24417	salmonella
314	1	256	0.7	6	0.7	387	1	PXBA_CANBO	Q00316	candida boi	387	6	0.7	297	1	VRPR_SALTY	P13041	salmonella
315	1	256	0.7	6	0.7	388	1	PXBB_CANBO	Q00317	candida boi	388	6	0.7	298	1	MOX2_XENLA	P39021	xenopus lae
316	1	257	0.7	6	0.7	389	1	HS70_LUPPO	P16121	lupinus pol	389	6	0.7	303	1	Y740_RICPR	O05961	rickettsia
317	1	257	0.7	6	0.7	390	1	LRR3_HUMAN	Q9by71	homo sapien	390	6	0.7	305	1	SGAU_MYCPN	P75294	mycoplasma
318	1	258	0.7	6	0.7	391	1	FABI_BACSU	P54616	bacillus su	391	6	0.7	306	1	FSLI_MOUSE	Q62356	mus musculus
319	1	258	0.7	6	0.7	392	1	GRAM_RAT	Q03238	rattus norv	392	6	0.7	307	1	FSLI_RAT	Q62632	rattus norv
320	1	259	0.7	6	0.7	393	1	IF2C_GALSU	Q08810	galdieria s	393	6	0.7	307	1	YJM6_YEAST	P47016	saccharomyc
321	1	261	0.7	6	0.7	394	1	HIS6_AZOBR	P26721	azospirilla	394	6	0.7	308	1	FSL1_HUMAN	Q12841	homo sapien
322	1	262	0.7	6	0.7	395	1	ATP6_HAEIN	P43719	haemophilus	395	6	0.7	308	1	FSL1_MACFA	Q9gky0	macaca fasc
323	1	262	0.7	6	0.7	396	1	FDHD_BACSU	P39756	bacillus su	396	6	0.7	308	1	VPT_HAEIN	P44227	haemophilus
324	1	262	0.7	6	0.7	397	1	TRUA_PYRHO	O58941	pyrococcus	397	6	0.7	308	1	Y880_METJA	O58290	methanococc
325	1	263	0.7	6	0.7	398	1	TRUA_BACHD	Q9z9j0	bacillus ha	398	6	0.7	310	1	YCUY_ECOLI	P76049	escherichia

107	7	0.9	1203	1	XCPE_XENLA	P50533 xenopus lae	180	6	0.7	155	1	IL17_HUMAN	Q16552 homo sapien
108	7	0.9	1209	1	THR_DROME	Q42286 drosophila	181	6	0.7	156	1	CYNS_PSEAE	Q91263 pseudomonas
109	7	0.9	1256	1	FLIL_DROME	Q24020 drosophila	182	6	0.7	157	1	END7_BPT4	P13340 bacterioph
110	7	0.9	1314	1	SWIL_YEAST	P09547 saccharomyc	183	6	0.7	158	1	LECG_TRIST	Q9ygl trimeresuru
111	7	0.9	1378	1	RPOB_MOUSE	Q82190 mus musculu	184	6	0.7	159	1	POP6_YEAST	P51275 pseudomonas
112	7	0.9	1379	1	RPOB_ODOSI	P49466 odontella s	185	6	0.7	160	1	ALGO_PSEAE	P32165 escherichia
113	7	0.9	1400	1	RON_HUMAN	Q04912 homo sapien	186	6	0.7	161	1	MENG_ECOLI	P11291 escherichia
114	7	0.9	1592	1	YNY2_YEAST	P53855 saccharomyc	187	6	0.7	161	1	YZCX_ECOLI	Q9zms3 helicobacte
115	7	0.9	1967	1	RRPO_PVNR	P17965 potato viru	188	6	0.7	162	1	AROK_HELPJ	P56073 helicobacte
116	7	0.9	2026	1	CYAA_YEAST	P08678 saccharomyc	189	6	0.7	162	1	AROK_HELPJ	P56073 helicobacte
117	7	0.9	2204	1	RRPL_NDV8	P11205 newcastlie d	190	6	0.7	162	1	WCAB_ECOLI	P77558 escherichia
118	7	0.9	2210	1	RRPO_TACV	P20430 tacaribe vi	191	6	0.7	163	1	YCBL_BACUN	P30906 bacteroides
119	7	0.9	3707	1	PGBM_MOUSE	Q05793 mus musculu	192	6	0.7	164	1	VAT_FWVD	P09521 figwort mos
120	7	0.9	4349	1	DYHC_FUSSO	P78716 fusarium so	193	6	0.7	166	1	THI2_BOVIN	Q95108 bos taurus
121	7	0.9	4367	1	DYHC_NEUCR	P45443 neurospora	194	6	0.7	166	1	THI2_HUMAN	Q99757 homo sapien
122	7	0.9	4393	1	PGBM_HUMAN	P98160 homo sapien	195	6	0.7	166	1	THI2_MOUSE	P97493 mus musculu
123	6	0.7	38	1	DPOB_BOVIN	Q27958 bos taurus	196	6	0.7	166	1	THI2_RAT	P97615 rattus norv
124	6	0.7	65	1	RL30_MYCTU	P35070 mycobacteri	197	6	0.7	169	1	COMD_METJA	P58415 methanococc
125	6	0.7	71	1	RL30_MYCLE	Q33001 mycobacteri	198	6	0.7	169	1	GBP_XENLA	Q93343 xenopus lae
126	6	0.7	72	1	YF77_HAEIN	Q57070 haemophilus	199	6	0.7	170	1	C550_SYNP2	Q55210 synechococc
127	6	0.7	75	1	YF67_THEMEA	Q9xlg3 thermotoga	200	6	0.7	172	1	NRDR_STRCL	O86848 streptomyce
128	6	0.7	77	1	GP46_BPSP1	O48400 bacterioph	201	6	0.7	176	1	IL7_BOVIN	P26895 bos taurus
129	6	0.7	78	1	YHGG_ECOLI	P46845 escherichia	202	6	0.7	176	1	IL7_SHEEP	Q28540 ovis aries
130	6	0.7	79	1	HM17_APIME	P15857 apis mellif	203	6	0.7	178	1	RIMM_PASMU	P37935 pasteurella
131	6	0.7	89	1	RS15_PSEPU	O87791 pseudomonas	204	6	0.7	180	1	GLUC_OCTDE	P22890 octodon deg
132	6	0.7	92	1	YG51_HAEIN	P44282 haemophilus	205	6	0.7	181	1	VG79_HSVI1	Q00148 ictaluriid h
133	6	0.7	98	1	GAS1_ARATH	P46689 arabidopsis	206	6	0.7	182	1	U131_ARATH	Q9m8t3 arabidopsis
134	6	0.7	98	1	NULM_RAT	P05507 rattus norv	207	6	0.7	182	1	YH25_STRCO	O59980 streptomyce
135	6	0.7	99	1	GAS2_ARATH	P46688 arabidopsis	208	6	0.7	183	1	AR21_CAEEL	Q9xwv3 caenorhabdi
136	6	0.7	99	1	GAS3_ARATH	P46687 arabidopsis	209	6	0.7	183	1	CORA_HPBV4	P03150 hepatitis b
137	6	0.7	104	1	SMT3_ARATH	P58552 arabidopsis	210	6	0.7	183	1	CORA_HPBVJ	P17391 hepatitis b
138	6	0.7	109	1	CYC6_CYACA	Q9tlw1 cyanidium c	211	6	0.7	183	1	CORA_HPBVJ	P12901 hepatitis b
139	6	0.7	109	1	VNST_BUNGE	P16992 bunyavirus	212	6	0.7	183	1	CORA_HPBVO	P17392 hepatitis b
140	6	0.7	113	1	FRT2_HUMAN	O75474 homo sapien	213	6	0.7	183	1	CORA_HPBVY	P03146 hepatitis b
141	6	0.7	113	1	Y109_HAEIN	P43943 haemophilus	214	6	0.7	183	1	CORA_HPBVZ	P03147 hepatitis b
142	6	0.7	118	1	RNPA_MYCPN	P75111 mycoplasma	215	6	0.7	183	1	RBS1_AMAHP	Q42516 amaranthus
143	6	0.7	118	1	YFBB_ECOLI	P24178 escherichia	216	6	0.7	185	1	CORA_HPBV2	P03149 hepatitis b
144	6	0.7	119	1	ACLY_ACHLY	P81730 achromobact	217	6	0.7	185	1	CORA_HPBVW	P03149 hepatitis b
145	6	0.7	120	1	RUS_ECOLI	P40116 escherichia	218	6	0.7	185	1	EFP_THEMEA	Q9x284 thermotoga
146	6	0.7	121	1	RK14_CYACA	Q9tlu2 cyanidium c	219	6	0.7	188	1	RIMM_DEIRA	Q9rsw1 deinococcus
147	6	0.7	123	1	RS13_MYCGE	P47421 mycoplasma	220	6	0.7	191	1	BID2_HAEIN	P45248 haemophilus
148	6	0.7	124	1	YAF2_DROME	Q9vni3 drosophila	221	6	0.7	191	1	CC42_CANAL	O4426 candida alb
149	6	0.7	126	1	Y4SK_RHLSN	P55654 rhizobium s	222	6	0.7	191	1	MOBA_CAMJE	Q9pmu9 campylobact
150	6	0.7	129	1	Y202_METJA	Q60267 methanococc	223	6	0.7	191	1	TNR6_ENTFC	Q06237 enterococcu
151	6	0.7	130	1	ATPE_MYCGA	P33255 mycoplasma	224	6	0.7	192	1	RR42_CYACA	O20229 cyanidium c
152	6	0.7	131	1	SODN_STRSO	P80734 streptomyce	225	6	0.7	193	1	NADD_BORBU	O51723 borrelia bu
153	6	0.7	134	1	ATPE_RHIME	Q92lk9 rhizobium m	226	6	0.7	194	1	GDL_DROME	P22468 drosophila
154	6	0.7	135	1	Y186_HAEIN	P44558 haemophilus	227	6	0.7	195	1	CORA_HPBVF	P29178 hepatitis b
155	6	0.7	136	1	ATPE_UREPA	Q9pr16 ureaplasma	228	6	0.7	195	1	FOS_SHEEP	O02761 ovis aries
156	6	0.7	138	1	Y076_MYCPN	P75555 mycoplasma	229	6	0.7	195	1	HIS7_WHEAT	P34048 triticum ae
157	6	0.7	139	1	IGF_MXGL	P22618 myxine glut	230	6	0.7	196	1	Y208_MYCGE	P47450 mycoplasma
158	6	0.7	140	1	Y128_LISMO	P58702 listeria mo	231	6	0.7	196	1	Y208_MYCPN	P75486 mycoplasma
159	6	0.7	140	1	Y175_LISIN	Q92fd1 listeria in	232	6	0.7	197	1	Y286_MYCPN	P75379 mycoplasma
160	6	0.7	142	1	PER_THEAC	P00218 thermoplas	233	6	0.7	197	1	YDB6_YEAST	Q12055 saccharomyc
161	6	0.7	142	1	PAL1_AGABI	Q92195 agaricus bi	234	6	0.7	201	1	CLP1_RHIL0	Q98m38 rhizobium l
162	6	0.7	142	1	RL13_PYRAB	Q9v196 pyrococcus	235	6	0.7	205	1	HIS7_PHARH	O94153 phaffia rho
163	6	0.7	143	1	ADXL_CHICK	P32140 gallus gall	236	6	0.7	205	1	PSB3_YEAST	P25451 saccharomyc
164	6	0.7	143	1	IL3_MACMU	P25140 macaca mula	237	6	0.7	205	1	RL3A_LUPLU	O49885 lupinus lut
165	6	0.7	143	1	MUCR_RHLSN	P55363 rhizobium s	238	6	0.7	206	1	RL3A_PICMA	O65055 picea maria
166	6	0.7	144	1	VMTT_LAMB	P03735 bacterioph	239	6	0.7	206	1	RIAA_ARATH	Q9sful arabidopsis
167	6	0.7	147	1	RL15_THEMEA	Q9x1j0 thermotoga	240	6	0.7	206	1	RIAB_ARATH	Q91rx8 arabidopsis
168	6	0.7	148	1	PTRI_METJA	Q27615 methanococc	241	6	0.7	206	1	RIAC_ARATH	Q9svr0 arabidopsis
169	6	0.7	148	1	RL9_THETH	P27151 thermus aqu	242	6	0.7	206	1	RIAD_ARATH	Q9fk0 arabidopsis
170	6	0.7	149	1	FUR_BACSU	P54574 bacillus su	243	6	0.7	207	1	DSBA_ERWCA	Q9rb10 erwinia car
171	6	0.7	149	1	Y213_METJA	Q57666 methanococc	244	6	0.7	207	1	DSBA_SALTI	O8x600 salmonella
172	6	0.7	150	1	VPL_BPHF1	P51722 bacterioph	245	6	0.7	207	1	DSBA_YERPE	O30848 salmonella
173	6	0.7	151	1	YCPR_VIBCH	P29483 vibrio chol	246	6	0.7	207	1	DSBA_YERPE	Q9xbv2 yersinia pe
174	6	0.7	151	1	Y128_AQAE	O66419 aquifex aeo	247	6	0.7	207	1	Y026_CAMJE	Q9pj85 campylobact
175	6	0.7	152	1	EXBB_PASHA	P72202 pasteurella	248	6	0.7	207	1	Y026_CAMJE	P54183 bacillus su
176	6	0.7	152	1	IL3_HUMAN	P08700 homo sapien	249	6	0.7	208	1	YCPD_BACSU	O35022 bacillus su
177	6	0.7	152	1	IL3_HYLLA	P06740 hylobates l	250	6	0.7	208	1	VATE_CHLMU	Q9pk83 chlamydia m
178	6	0.7	152	1	IL3_PANTR	P28809 pan troglod	251	6	0.7	210	1	HIS1_SYNF3	Q55503 synechocyst
179	6	0.7	154	1	PTGA_MYCCA	P45618 mycoplasma	252	6	0.7	211	1	COBL_METJA	Q58917 methanococc

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2002, 09:44:53 ; Search time 15.77 Seconds  
(without alignments)  
1981.400 Million cell updates/sec

Title: US-09-202-054-1\_COPY\_30\_836  
Perfect score: 807  
Sequence: 1 FPKTLPDVTLDVPRNHVIV.....PGAHKGQSVISLDLYTCELD 807

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 6

Total number of hits satisfying chosen parameters: 1410

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	807	100.0	1049	1	TLR7_HUMAN
2	34	4.2	1050	1	TLR7_MOUSE
3	17	2.1	1032	1	TLR8_MOUSE
4	10	1.2	1041	1	TLR8_HUMAN
5	9	1.1	359	1	PGS2_HUMAN
6	9	1.1	360	1	PGS2_BOVIN
7	9	1.1	360	1	PGS2_CANFA
8	9	1.1	360	1	PGS2_PIG
9	9	1.1	360	1	PGS2_RABIT
10	9	1.1	626	1	GPBA_HUMAN
11	9	1.1	826	1	TLR4_PAPAN
12	9	1.1	839	1	TLR4_HUMAN
13	9	1.1	839	1	TLR4_PANPA
14	9	1.1	1269	1	FLIH_HUMAN
15	9	1.1	1315	1	CHAO_DROME
16	9	1.1	1770	1	RIL5_YEAST
17	8	1.0	274	1	VGLL_MCMVK
18	8	1.0	274	1	VGLL_MCMVS
19	8	1.0	283	1	YORS_ADEGI
20	8	1.0	288	1	PDXK_SALTY
21	8	1.0	796	1	TLR6_HUMAN
22	8	1.0	806	1	BIMA_EMENI
23	8	1.0	1042	1	SYI_BORBU
24	8	1.0	1433	1	CATE_YEAST
25	7	0.9	41	1	RL36_RICPR
26	7	0.9	91	1	NULM_BRALA
27	7	0.9	117	1	ARR1_ECOLI
28	7	0.9	124	1	YG3E_YEAST
29	7	0.9	125	1	VG61_BPMU5
30	7	0.9	155	1	RIMM_RICPR
31	7	0.9	177	1	RELX_MESAU
32	7	0.9	180	1	YC47_METJA
33	7	0.9	196	1	Y628_CHLMU

197	1	TSAA_BUCAI	197	1	P52729	buchnera sp
198	1	VP24_NPVAC	198	1	P41678	autographa
212	7	ERP5_YEAST	212	7	P38819	saccharomyc
222	7	RPOA_RECAM	222	7	O21260	reclinomona
224	7	DEOC_MYCPN	224	7	P09924	mycoplasma
237	7	YMI17_YEAST	237	7	P40207	saccharomyc
266	7	CB21_PINTH	266	7	P10049	pinus thunb
274	7	PERA_ECO27	274	7	P43459	escherichia
318	7	Y692_METH	318	7	O26788	methanobact
319	7	YDDK_ECOLI	319	7	P76123	escherichia
319	7	ERMS_STRFR	319	7	P45439	streptomyce
323	7	YECF_ECOLI	323	7	P76291	escherichia
327	7	MDH_THEFL	327	7	P10584	thermus aqu
333	7	PARE_MYCLE	333	7	O50201	mycobacteri
338	7	RTCA_METJA	338	7	O60335	methanococc
341	7	KITH_VZVD	341	7	P09250	varicella-z
341	7	T4BB_BACCO	341	7	O07606	bacillus co
345	7	ARGC_BACHD	345	7	O9k8v2	bacillus ha
352	7	KERA_BOVIN	352	7	O62702	bos taurus
352	7	KERA_HUMAN	352	7	O60938	homo sapien
353	7	KERA_CHICK	353	7	O42235	gallus gall
353	7	KERA_COTJA	353	7	O9de66	coturnix co
369	7	WNT1_AMBME	369	7	P21551	ambystoma m
370	7	WNT1_HUMAN	370	7	P04628	homo sapien
370	7	WNT1_MOUSE	370	7	P04426	mus musculu
371	7	WNT1_XENLA	371	7	P10108	xenopus lae
387	7	SASA_SYNP7	387	7	O06804	synecococc
395	7	ARP2_DROME	395	7	P45888	drosophila
420	7	DCDA_ECOLI	420	7	F00861	escherichia
429	7	MTAL_ACEPA	429	7	O52702	acetobacter
430	7	SERC_ARATH	430	7	O96255	arabidopsis
449	7	COMB_STRPN	449	7	P36498	streptococc
456	7	YGL3_ASTLO	456	7	P14761	astasia lon
462	7	BIND_LITVA	462	7	P23117	lytechinus
477	7	UL49_HSV7J	477	7	P52442	human herpe
504	7	IC1_MOUSE	504	7	P97290	mus musculu
508	7	IC1_MEDTR	508	7	O9znx6	medicago tr
511	7	DDC_HAEIN	511	7	P71362	haemophilus
512	7	C6D2_DROME	512	7	O9w223	drosophila
512	7	CP11_HUMAN	512	7	P04798	homo sapien
512	7	CP11_MACFA	512	7	P33616	macaca fasc
528	7	LADI_MOUSE	528	7	P57016	mus musculu
549	7	PPQ1_YEAST	549	7	P32945	saccharomyc
564	7	ARAB_BACST	564	7	O9s468	bacillus st
568	7	DPOL_BPCP1	568	7	O37989	bacterioph
582	7	SHO2_HUMAN	582	7	O9uq13	homo sapien
582	7	SHO2_MOUSE	582	7	O88520	mus musculu
590	7	YNNW_YEAST	590	7	P53863	saccharomyc
598	7	TDPH_CAEEL	598	7	O22703	caenorhabdi
631	7	IL16_HUMAN	631	7	O14005	homo sapien
657	7	HUTH_HUMAN	657	7	P42357	homo sapien
659	7	FPS1_YEAST	659	7	P23900	saccharomyc
676	7	HS7C_TRYBB	676	7	P20030	trypanosoma
713	7	GAC1_HUMAN	713	7	O75325	homo sapien
732	7	ACPH_HUMAN	732	7	P13798	homo sapien
732	7	ACPH_PIG	732	7	P19205	sus scrofa
732	7	ACPH_RAT	732	7	P13676	rattus norv
733	7	YFJ2_YEAST	733	7	P43602	saccharomyc
734	7	GL3B_DICDI	734	7	P34116	dictyostell
811	7	TLR4_HUMAN	811	7	O9bxf5	homo sapien
830	7	YJG2_YEAST	830	7	P40367	saccharomyc
835	7	TLR4_RAT	835	7	O9qx05	rattus norv
838	7	TLR4_CRIGR	838	7	O9w482	cricetulus
859	7	TLR5_MOUSE	859	7	O9j1f7	mus musculu
880	7	LYTD_BACSU	880	7	P39848	bacillus su
885	7	YDGH_BACSU	885	7	P96706	bacillus su
904	7	TLR3_HUMAN	904	7	O15455	homo sapien
959	7	RLK5_ARATH	959	7	P47735	arabidopsis
1025	7	ADAL_YEAST	1025	7	P38065	saccharomyc
1032	7	TLR9_MOUSE	1032	7	O9eqf3	mus musculu
104	7	MGCI_HUMAN	104	7	O60732	homo sapien
105	7	CD45_MOUSE	105	7	P06800	mus musculu
106	7	SCII_CHICK	106	7	O90988	gallus gall



FT SIGNAL 1 16 POTENTIAL.  
FT PROPEP 17 30 BY SIMILARITY.  
FT CHAIN 31 360 BONE PROTEOGLYCAN II.  
FT REPEAT 78 99 LRR 1.  
FT REPEAT 100 123 LRR 2.  
FT REPEAT 124 146 LRR 3.  
FT REPEAT 147 168 LRR 4.  
FT REPEAT 169 194 LRR 5.  
FT REPEAT 195 218 LRR 6.  
FT REPEAT 219 239 LRR 7.  
FT REPEAT 240 263 LRR 8.  
FT REPEAT 264 286 LRR 9.  
FT REPEAT 287 309 LRR 10.  
FT CARBOHYD 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
SMILARITY).  
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 360 AA; 39896 MW; 0B50C6756FE02369 CRC64;  
Query Match 1.1%; Score 9; DB 1; Length 360;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 113 IPQGLPPSL 121  
|||||||  
Db 216 IPQGLPPSL 224  
RESULT 10  
GPBA\_HUMAN  
ID GPBA\_HUMAN STANDARD; PRT; 626 AA.  
AC P07359;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Platelet glycoprotein IB alpha chain precursor (GP-IB alpha) (GP1BA)  
DE (CD42B-alpha) (CD42B) [Contains: Glycocalicin].  
GN GP1BA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=87289655; PubMed=3303030;  
RA Lopez J.A., Chung D.W., Fujikawa K., Hagen F.S., Papayannopoulou T.,  
Roth G.J.;  
RT "Cloning of the alpha chain of human platelet glycoprotein Ib: a  
transmembrane protein with homology to leucine-rich alpha 2-  
glycoprotein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89025874; PubMed=2845978;  
RA Wenger R.H., Kieffer N., Wicki A.N., Clemetson K.J.;  
RT "Structure of the human blood platelet membrane glycoprotein Ib alpha  
gene.";  
RL Biochem. Biophys. Res. Commun. 156:389-395(1988).  
RN [3]  
RP SEQUENCE OF 17-315.  
RX MEDLINE=87289654; PubMed=3497398;  
RA Titani K., Takio K., Handa M., Ruggeri Z.M.;  
RT "Amino acid sequence of the von Willebrand factor-binding domain of  
platelet membrane glycoprotein Ib.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614(1987).  
RN [4]  
RP DISULFIDE BONDS.  
RX MEDLINE=91301149; PubMed=2070794;  
RA Hess D., Schaller J., Rickli E.E., Clemetson K.J.;  
RT "Identification of the disulphide bonds in human platelet  
glycocalicin.";

RL Eur. J. Biochem. 199;389-393(1991).  
RN [5]  
RP VARIANT S1BA.  
RX MEDLINE=92265982; PubMed=1586750;  
RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,  
Ruggeri Z.M.;  
RT "Genetic and structural characterization of an amino acid dimorphism  
in glycoprotein Ib alpha involved in platelet transfusion  
refractoriness.";  
RL Blood 79:3086-3090(1992).  
RN [6]  
RP VARIANT BSS PHE-73.  
RX MEDLINE=92110577; PubMed=1730088;  
RA Miller J.L., Lyle V.A., Cunningham D.;  
RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein  
Ib alpha leucine tandem repeat occurring in patients with an  
autosomal dominant variant of Bernard-Soulier disease.";  
RL Blood 79:439-446(1992).  
RN [7]  
RP VARIANT BSS VAL-172.  
RX MEDLINE=93388851; PubMed=7690774;  
RA Ware J., Russell S.R., Marchese P., Murata M., Mazucato M.,  
de Marco L., Ruggeri Z.M.;  
RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib  
alpha resulting in the Bernard-Soulier syndrome.";  
RL J. Clin. Invest. 92:1213-1220(1993).  
RN [8]  
RP VARIANT BSS SER-225.  
RX MEDLINE=95118882; PubMed=7819107;  
RA Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,  
Ribera A., Gallardo D.;  
RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha  
gene is associated with Bernard-Soulier syndrome.";  
RL Br. J. Haematol. 88:839-844(1994).  
RN [9]  
RP VARIANT PSEUDO-VWD VAL-249.  
RX MEDLINE=91271273; PubMed=2052556;  
RA Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;  
RT "Mutation in the gene encoding the alpha chain of platelet  
glycoprotein Ib in platelet-type von Willebrand disease.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:4761-4765(1991).  
RN [10]  
RP VARIANT PSEUDO-VWD VAL-249.  
RX MEDLINE=93253059; PubMed=8486780;  
RA Murata M., Russell S.R., Ruggeri Z.M., Ware J.;  
RT "Expression of the phenotypic abnormality of platelet-type von  
Willebrand disease in a recombinant glycoprotein Ib alpha fragment.";  
RL J. Clin. Invest. 91:2133-2137(1993).  
RN [11]  
RP VARIANT PSEUDO-VWD VAL-255.  
RX MEDLINE=93214031; PubMed=8384898;  
RA Russell S.D., Roth G.J.;  
RT "Pseudo-von Willebrand disease: a mutation in the platelet  
glycoprotein Ib alpha gene associated with a hyperactive surface  
receptor.";  
RL Blood 81:1787-1791(1993).  
RN [12]  
RP VARIANT BSS LEU-195 DEL.  
RX MEDLINE=95178321; PubMed=7873390;  
RA de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,  
Chevallier J., Gachet C., Briquel M.-E., Cazenave J.-P.;  
RT "A three-base deletion removing a leucine residue in a leucine-rich  
repeat of platelet glycoprotein Ib alpha associated with a variant of  
Bernard-Soulier syndrome (Nancy I).";  
RL Br. J. Haematol. 89:386-396(1995).  
CC -!- FUNCTION: GP-IB, A SURFACE MEMBRANE PROTEIN OF PLATELETS,  
PARTICIPATES IN THE FORMATION OF PLATELET PLUGS BY BINDING TO VON  
WILLEBRAND FACTOR, WHICH IS ALREADY BOUND TO THE SUBENDOTHELIUM.  
CC -!- SUBUNIT: GP-IB ALPHA AND BETA ARE DISULFIDE LINKED. GP-IX IS  
CC -!- COMPLEXED WITH THE GP-IB HETERODIMER VIA A NON COVALENT LINKAGE.  
CC -!- PTM: GLYCOCALICIN, WHICH IS APPROXIMATELY COEXTENSIVE WITH THE  
EXTRACELLULAR PART OF THE MOLECULE, IS CLEAVED OFF BY CALPAIN

CC DURING PLATELET LYSIS.  
CC -!- POLYMORPHISM: POSITION 161 IS ASSOCIATED WITH PLATELET-SPECIFIC  
CC ALLANTHIN SIBA. SIBA(-) HAS THR-161 AND SINA(+) HAS MET-161.  
CC SIBA IS INVOLVED IN NEONATAL ALLOIMMUNE THROMBOCYTOPENIA (NATP).  
CC -!- DISEASE: DEFECTS IN GPIIA ARE ONE OF THE CAUSES OF BERNARD-SOULIER  
CC SYNDROME (BSS). BSS PATIENTS HAVE UNUSUALLY LARGE PLATELETS AND  
CC HAVE A CLINICAL BLEEDING TENDENCY.  
CC -!- DISEASE: DEFECTS IN GPIIA ARE ONE OF THE CAUSES OF VON WILLEBRAND  
CC DISEASE (VWD) KNOWN AS PLATELET-TYPE VON WILLEBRAND DISEASE OR  
CC PSEUDO-VON WILLEBRAND DISEASE. THIS AUTOSOMAL DOMINANT BLEEDING  
CC DISORDER IS CAUSED BY AN INCREASED AFFINITY OF GP-IB FOR SOLUBLE  
CC VWF RESULTING IN IMPAIRED HEMOSTATIC FUNCTION DUE TO THE REMOVAL  
CC OF VWF FROM THE CIRCULATION.  
CC -!- MISCELLANEOUS: PLATELET ACTIVATION APPARENTLY INVOLVES DISRUPTION  
CC OF THE MACROMOLECULAR COMPLEX OF GP-IB WITH THE PLATELET  
CC GLYCOPROTEIN IX (GP-IX) AND DISSOCIATION OF GP-IB FROM THE  
CC ACTIN-BINDING PROTEIN.  
CC -!- MISCELLANEOUS: BINDING SITES FOR VON WILLEBRAND FACTOR AND  
CC THROMBIN (THE LATTER SITE WITH UNKNOWN FUNCTION) ARE IN THE  
CC AMINO-TERMINAL PART OF THE MOLECULE.  
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J02940; AAA52595.1; -;  
DR EMBL: M22403; AAA52596.1; -;  
DR PIR: A27075; NBHU1A.  
DR GlycoSuiteDB: P07359; -;  
DR MIM: 231200; -;  
DR MIM: 177820; -;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR\_Cterm.  
DR InterPro: IPR000372; LRR\_Nterm.  
DR InterPro: IPR003592; LRR\_Out.  
DR InterPro: IPR003591; LRR\_typ.  
DR Pfam: PF00560; LRR: 6.  
DR Pfam: PF01463; LRRCT: 1.  
DR Pfam: PF01462; LRRNT: 1.  
DR PRINTS: PR00019; LEURICHRPT.  
DR SMART: SM00370; LRR: 3.  
DR SMART: SM00082; LRRCT: 1.  
DR SMART: SM00013; LRRNT: 1.  
DR SMART: SM00369; LRR\_typ: 2.  
DR Platelet; Transmembrane; Glycoprotein; Hemostasis; Blood coagulation;  
KW Repeat; Leucine-rich repeat; Signal; Cell adhesion; Disease mutation;  
KW Polymorphism; von Willebrand disease; Bernard Soulier syndrome.  
FT SIGNAL 1 16  
FT CHAIN 17 626 PLATELET GLYCOPROTEIN IB ALPHA CHAIN.  
FT CHAIN 17 ?  
FT DOMAIN 17 505 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 526 526 POTENTIAL.  
FT DOMAIN 527 626 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 70 92 LRR 1.  
FT REPEAT 93 117 LRR 2.  
FT REPEAT 119 138 LRR 3.  
FT REPEAT 139 162 LRR 4.  
FT REPEAT 164 186 LRR 5.  
FT REPEAT 188 210 LRR 6.  
FT REPEAT 379 396 THR/PRO-RICH.  
FT REPEAT 387 395 THR/PRO-RICH.  
FT REPEAT 400 408 THR/PRO-RICH.  
FT REPEAT 409 417 THR/PRO-RICH.  
FT REPEAT 422 430 THR/PRO-RICH.  
FT DISULFID 20 33  
FT DISULFID 225 264  
FT DISULFID 227 280  
FT CARBOHYD 37  
N-LINKED (GLCNAC. . .).

FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 308 308 O-LINKED.  
FT VARIANT 72 72 R -> H (IN DBSNP:6068).  
FT /FTID-VAR\_011909.  
FT VARIANT 73 73 L -> F (IN BSS).  
FT /FTID-VAR\_005256.  
FT VARIANT 161 161 T -> M (IN ALLOANTIGEN SIBA(+)).  
FT /FTID-VAR\_005257.  
FT VARIANT 172 172 A -> V (IN BSS).  
FT /FTID-VAR\_005258.  
FT VARIANT 195 195 MISSING (IN BSS).  
FT /FTID-VAR\_005259.  
FT VARIANT 225 225 C -> S (IN BSS).  
Query Match 1.1% Score 9; DB 1; Length 626;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 674 TLDLSHNLQ 682  
|||||  
DB 97 TLDLSHNLQ 105  
RESULT 11  
TLR4\_PAPAN  
ID TLR4\_PAPAN STANDARD; PRT; 826 AA.  
AC Q9TSP2;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DE Toll-like receptor 4 precursor.  
GN TLR4.  
OS Papio anubis (Olive baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=11104518;  
RA Smirnova I., Poltorak A., Chan E.K.L., McBride C., Beutler B.:  
RT "Phylogenetic variation and polymorphism at the Toll-like receptor 4  
RL (In) Genome Biol. 1:RESEARCH002.1-2.10(2000).  
CC -!- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate  
CC immune response to bacterial lipopolysaccharide (LPS). Acts via  
CC MyD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine  
CC secretion and the inflammatory response (By similarity).  
CC -!- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a  
CC multi-protein complex containing at least CD14, MD-2 and TLR4.  
CC Binds MD-2 via the extracellular domain. Binds MyD88 and TIRAP via  
CC their respective TIR domains (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF180964; AAF07059.1; -;  
DR EMBL: AF180962; AAF07059.1; JOINED.  
DR EMBL: AF180963; AAF07059.1; JOINED.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR\_Cterm.  
DR InterPro: IPR003592; LRR\_Out.  
DR InterPro: IPR003591; LRR\_typ.  
DR InterPro: IPR000157; TIR.

DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; TIR; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 1.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYD; 2.  
DR SMART; SM00255; TIR; 1.  
DR PROSITE; PS0104; TIR; 1.  
KW Receptor; Immune response; Inflammatory response; Signal;  
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.  
FT SIGNAL 1 23  
FT CHAIN 24 826  
FT DOMAIN 24 631  
FT TRANSMEM 632 652  
FT DOMAIN 653 826  
FT REPEAT 53 76  
FT REPEAT 77 100  
FT REPEAT 101 124  
FT REPEAT 128 149  
FT REPEAT 150 173  
FT REPEAT 174 197  
FT REPEAT 203 225  
FT REPEAT 228 252  
FT REPEAT 277 303  
FT REPEAT 327 350  
FT REPEAT 351 372  
FT REPEAT 373 398  
FT REPEAT 400 421  
FT REPEAT 422 445  
FT REPEAT 447 469  
FT REPEAT 470 494  
FT REPEAT 495 518  
FT REPEAT 520 541  
FT REPEAT 543 569  
FT REPEAT 571 592  
FT DOMAIN 572 818  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 826 AA; 94678 MW; 42277318E5F1769 CRC64;

Query Match 1.1%; Score 9; DB 1; Length 826;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 723 LDLSNKKIQ 731  
|||||||  
Db 180 LDLSNKKIQ 188

RESULT 12  
TLR4\_HUMAN  
ID TLR4\_HUMAN STANDARD; PRT; 839 AA.  
AC 000206; Q9UK78; Q9UM57;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Toll-like receptor 4 precursor (hToll).  
GN TLR4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

SEQUENCE FROM N.A.  
RP TISSUE-Spleen;  
RX MEDLINE=97379437; PubMed=9237759;  
RA Medzhitov R., Preston-Hurlburt P., Janeway C.A. Jr.;  
RT "A human homologue of the Drosophila Toll protein signals activation  
of adaptive immunity.";  
RL Nature 388:394-397(1997).  
RN [2]  
RP SEQUENCE OF 41-839 FROM N.A.  
RC TISSUE-Lung, Placenta, and Fetal liver;  
RX MEDLINE=98118556; PubMed=9435236;  
RA Rock F.L., Hardman G., Timans J.C., Kastelein R.A., Bazan J.F.;  
RT "A family of human receptors structurally related to Drosophila  
Toll.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANTS GLY-299 AND ILE-399.  
RX PubMed=11104518;  
RA Smirnova I., Poitorak A., Chan E.K.L., McBride C., Beutler B.;  
RT "Phylogenetic variation and polymorphism at the Toll-like receptor 4  
locus (TLR4).";  
RL (In) Genome Biol. 1:RESEARCH002.1-2.10(2000).  
RN [4]  
RP SEQUENCE FROM N.A., AND VARIANTS GLY-299 AND ILE-399.  
RX PubMed=10835634;  
RA Arbour N.C., Lorenz E., Schutte B.C., Zabner J., Kline J.N., Jones M.,  
RT "TLR4 mutations are associated with endotoxin hyporesponsiveness in  
humans.";  
RL Nat. Genet. 25:187-191(2000).  
RN [5]  
RP MUTAGENESIS OF GLU-697; ARG-710; ASP-711 AND PRO-714.  
RX PubMed=11081518;  
RA Xu Y., Tao X., Shen B., Horng T., Medzhitov R., Manley J.L., Tong L.;  
RT "Structural basis for signal transduction by the Toll/interleukin-1  
receptor domains.";  
RL Nature 408:111-115(2000).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND MUTAGENESIS OF ASN-526 AND ASN-575.  
RX PubMed=11706042;  
RA da Silva Correia J., Ulevitch R.J.;  
RT "MD-2 and TLR4 N-linked glycosylations are important for a functional  
lipopolysaccharide receptor.";  
RL J. Biol. Chem. 277:1845-1854(2002).  
CC -!- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate  
immune response to bacterial lipopolysaccharide (LPS). Acts via  
MyD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine  
secretion and the inflammatory response.  
CC -!- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a  
multi-protein complex containing at least CD14, MD-2 and TLR4.  
CC Binds MD-2 via the extracellular domain. Binds MyD88 and TIRAP via  
their respective TIR domains.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Highly expressed in placenta, spleen and  
peripheral blood leukocytes. Detected in monocytes, macrophages,  
dendritic cells and several types of T-cells.  
CC -!- PTM: N-glycosylated. Glycosylation of Asn-526 and Asn-575 seems to  
be necessary for the expression of TLR4 on the cell surface and  
the LPS-response. Likewise, mutants lacking two or more of the  
other N-glycosylation sites were deficient in interaction with  
LPS.  
CC -!- POLYMORPHISM: Allele B (Gly-299, Ile-399) is associated with a  
blunted response to inhaled LPS.  
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -!- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)





```

Db 109 LDLSHNQLT 117
|||||||
RESULT 15
ID CHAO_DROME STANDARD; PRT: 1315 AA.
AC P12024; OSVA01;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Chaoptin precursor (Photoreceptor cell-specific membrane protein).
GN CHP OR CHT OR CGI744.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RA MEDLINE=88135762; PubMed=3124963;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Duerbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Stimpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
RP SEQUENCE OF 30-50 FROM N.A.
RC TISSUE=Head;
RX MEDLINE=85166231; PubMed=3920657;

```

```

RA Zipursky S.L., Venkatesh T.R., Benzer S.;
RT "From monoclonal antibody to gene for a neuron-specific glycoprotein
in Drosophila.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:1855-1859(1985).
[4]
RP SEQUENCE OF 30-50, AND TISSUE SPECIFICITY.
RC STRAIN=CANTON-S; TISSUE=Head;
RX MEDLINE=84106810; PubMed=6420071;
RA Zipursky S.L., Venkatesh T.R., Teplow D.B., Benzer S.;
RT "Neuronal development in the Drosophila retina: monoclonal antibodies
as molecular probes.";
RL Cell 36:15-26(1984).
CC -|- FUNCTION: REQUIRED FOR PHOTORECEPTOR CELL MORPHOGENESIS. MEDIATES
CC HOMOPHILIC CELLULAR ADHESION
CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR SURFACE OF R-CELL PLASMA
CC MEMBRANE.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN PHOTORECEPTOR CELLS AND THEIR
CC AXONS IN THE ADULT RETINA. THE OCELLUS AND LARVAL PHOTORECEPTOR
CC ORGAN.
CC -|- DEVELOPMENTAL STAGE: EXPRESSED 24 HOURS AFTER INITIATION OF
CC PHOTORECEPTOR CELL DIFFERENTIATION, PERSISTS THROUGH DEVELOPMENT
CC TO ADULTHOOD.
CC -|- SIMILARITY: BELONGS TO THE CHAOPTIN FAMILY.
CC -|- SIMILARITY: CONTAINS 38 LEUCINE-RICH REPEATS (LRR).
CC -|- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 1123.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; M19017; AAA28425.1; ALT_FRAME.
CC EMBL; M19008; AAA28425.1; JOINED.
CC EMBL; M19009; AAA28425.1; JOINED.
CC EMBL; M19010; AAA28425.1; JOINED.
CC EMBL; M19011; AAA28425.1; JOINED.
CC EMBL; M19012; AAA28425.1; JOINED.
CC EMBL; M19013; AAA28425.1; JOINED.
CC EMBL; M19014; AAA28425.1; JOINED.
CC EMBL; M19016; AAA28425.1; JOINED.
CC EMBL; AE003777; AA57127.1; -.
CC EMBL; K03274; AAA28651.1; -.
CC PIR; A29944; A29944.
CC FlyBase; FBgn0000313; chp.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003592; LRR_out.
CC InterPro; IPR003591; LRR_typ.
CC Pfam; PF00560; LRR; 29.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00370; LRR; 9.
CC SMART; SM00369; LRR_Typ; 27.
CC Glycoprotein; Membrane; Signal; Repeat; Leucine-rich repeat; Vision.
FT SIGNAL 1 29
FT CHAIN 30 1315 CHAOPTIN.
FT REPEAT 101 124 LRR 1.
FT REPEAT 126 149 LRR 2.
FT REPEAT 150 173 LRR 3.
FT REPEAT 175 198 LRR 4.
FT REPEAT 199 222 LRR 5.
FT REPEAT 224 247 LRR 6.
FT REPEAT 249 272 LRR 7.
FT REPEAT 277 300 LRR 8.
FT REPEAT 302 324 LRR 9.
FT REPEAT 326 347 LRR 10.
FT REPEAT 349 372 LRR 11.
FT REPEAT 374 397 LRR 12.
FT REPEAT 451 474 LRR 13.
FT REPEAT 475 498 LRR 14.
FT REPEAT 525 548 LRR 15.

```

```
FT REPEAT 550 572 LRR 16.
FT REPEAT 575 598 LRR 17.
FT REPEAT 599 622 LRR 18.
FT REPEAT 624 646 LRR 19.
FT REPEAT 648 670 LRR 20.
FT REPEAT 674 697 LRR 21.
FT REPEAT 706 729 LRR 22.
FT REPEAT 731 754 LRR 23.
FT REPEAT 755 778 LRR 24.
FT REPEAT 779 802 LRR 25.
FT REPEAT 803 826 LRR 26.
FT REPEAT 828 849 LRR 27.
FT REPEAT 852 875 LRR 28.
FT REPEAT 877 900 LRR 29.
FT REPEAT 902 924 LRR 30.
FT REPEAT 926 947 LRR 31.
FT REPEAT 948 970 LRR 32.
FT REPEAT 971 994 LRR 33.
FT REPEAT 995 1017 LRR 34.
FT REPEAT 1043 1068 LRR 35.
FT REPEAT 1117 1140 LRR 36.
FT REPEAT 1142 1165 LRR 37.
FT REPEAT 1169 1193 LRR 38.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 422 422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 718 718 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 936 936 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 970 970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1012 1012 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1122 1122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1152 1152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1171 1171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 12 12 A -> V (IN REF. 1).
FT CONFLICT 44 44 C -> H (IN REF. 4).
FT CONFLICT 50 50 C -> H (IN REF. 4).
FT CONFLICT 937 937 I -> V (IN REF. 1).
SQ SEQUENCE 1315 AA; 151982 MW; D2D89A64EB46FCE5 CRC64;

Query Match 1.1%; Score 9; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 222 LQILDLSGN 230
Db 352 LQILDLSGN 360
|||||
RESULT 16
RI15_YEAST STANDARD; PRT; 1770 AA.
AC P43565;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase RIM15 (EC 2.7.1.-).
GN RIM15 OR TAK1 OR YFL033C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
```

```
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae";
RL Nat. Genet. 10:261-268(1995).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97265402; PubMed=9111339;
RA Vidan S., Mitchell A.P.;
RT "Stimulation of yeast meiotic gene expression by the
RT glucose-repressible protein kinase Rim15p.";
RL Mol. Cell. Biol. 17:2688-2697(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Reinders A., Buerckert N., Boller T., Wiemken A., de Virgilio C.;
RT "Saccharomyces cerevisiae CAMP-dependent protein kinase entry
RT into stationary phase through the Rim15p protein kinase.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE ACTIVATION OF A MEIOTIC GENES ACTIVATION
CC PATHWAY.
CC -1- PTM: AUTOPHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- STRONG, WITH S.POMBE CEK1.
CC -1- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
EMBL; D50617; BAA09206.1; -
EMBL; U83459; AAB64088.1; -
EMBL; AJ001030; CAA04486.1; -
HSP; Q16539; IWFC.
SGD; S0001861; RIM15.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 2.
DR Pfam; PF00072; response_reg; 1.
DR SMART; SM00448; REC; 1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Meiosis.
FT DOMAIN 794 1254 PROTEIN KINASE.
FT DOMAIN 1636 1750 RESPONSE REGULATORY.
FT DOMAIN 343 358 POLY-ASN.
FT DOMAIN 620 624 POLY-SER.
FT NP_BIND 800 808 ATP (BY SIMILARITY).
FT BINDING 823 823 ATP (BY SIMILARITY).
FT ACT_SITE 918 918 BY SIMILARITY.
FT DOMAIN 975 980 POLY-ASN.
FT DOMAIN 1213 1218 POLY-GLU.
FT DOMAIN 1386 1391 POLY-THR.
SQ SEQUENCE 1770 AA; 196530 MW; DC1064825000FAFF CRC64;

Query Match 1.1%; Score 9; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 646 PNKLNLSLA 654
|||||
```

```
Db 1397 PNLKNLSLA 1405

RESULT 17
VGLL_MCMVK STANDARD; PRT; 274 AA.
AC P52513;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glycoprotein L precursor.
GN GL OR UL115.
OS Murine cytomegalovirus (strain KI81).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=69156;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95053910; PubMed=7964634;
RU Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,
RA Shellam G.R.;
RT "Identification, sequencing and expression of the glycoprotein L gene
of murine cytomegalovirus";
RL J. Gen. Virol. 75:3235-3240(1994).
CC -!- FUNCTION: ASSOCIATED WITH GLYCOPROTEIN H (GH) TO FORM A COMPLEX
CC IMPORTANT FOR INFECTION AND CELL FUSION. THIS ASSOCIATION IS
CC NECESSARY FOR THE CORRECT PROCESSING AND CELL SURFACE EXPRESSION
CC OF GH (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN L.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L32962; AAA57343.1; -
DR InterPro; IPR002689; Cytomegalo_gL.
DR Pfam; PF01801; Cytomegalo_gL; 1.
KW Glycoprotein; Signal; Envelope protein.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 274 GLYCOPROTEIN L.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 274 AA; 31239 MW; 504D349FDBEC5C03 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 274;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 505 NGSEFQPL 512
Db 64 NGSEFQPL 71
|||||||

RESULT 19
YOR5_ADEG1 STANDARD; PRT; 283 AA.
AC P20747;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 31.5 kDa protein (ORF 5) (ORF10).
OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX NCBI_TaxID=10553;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90251474; PubMed=21600072;
RU Akopian T.A., Kruglyak V.A., Rivkina M.B., Naroditsky B.S.,
RA Tikhonenko T.I.;
RT "Sequence of an avian adenovirus (CELO) DNA fragment (0-11.2%)."
RL Nucleic Acids Res. 18:2825-2825(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96186720; PubMed=8627769;
RU Chiocca S., Kurzbauer R., Schaffner G., Baker A., Mautner V.,
RA Cotten M.;
RT "The complete DNA sequence and genomic organization of the avian
adenovirus CELO."
RL J. Virol. 70:2939-2949(1996).
CC -----
```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X17217; CAA35087.1; -;  
DR EMBL; U46933; AAC54931.1; -;  
DR PIR; S10005; S10005.  
KW Hypothetical protein.  
SQ SEQUENCE 283 AA; 31487 MW; 6018412DA598183D CRC64;

Query Match 1.0%; Score 8; DB 1; Length 283;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 TNLTLTIN 46  
|||||||  
DB 222 TNLTLTIN 229

## RESULT 20

ID PDXK\_SALTY STANDARD; PRT; 288 AA.  
AC P40192;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Pyridoxine kinase (EC 2.7.1.35) (Pyridoxal kinase) (Vitamin B6  
kinase) (Pyridoxamine kinase) (PN/PL/PM kinase).  
GN PDXK OR SPM2435.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2;  
RX MEDLINE=95337418; PubMed=7612925;  
RA Titzgemyer F.M., Reizer J., Reizer A., Tang J., Parr T.R. Jr.,  
RA Sailer M.H. Jr.;  
RT "Nucleotide sequence of the region between crr and cysM in Salmonella  
typhimurium: five novel ORFs including one encoding a putative  
transcriptional regulator of the phosphotransferase system.";  
RL DNA Seq. 5:145-152(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
LT2.";  
RL Nature 413:852-856(2001).  
CC -!- FUNCTION: PHOSPHORYLATE B6 VITAMERS; FUNCTIONS IN A SALVAGE  
CC PATHWAY. USES PYRIDOXAL, PYRIDOXINE, AND PYRIDOXAMINE AS  
CC SUBSTRATES.  
CC -!- CATALYTIC ACTIVITY: ATP + pyridoxal = ADP + pyridoxal 5'-  
CC phosphate.  
CC -!- COPACITOR: ZINC OR MAGNESIUM (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE PYRIDOXINE KINASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U11243; AAC43343.1; -;  
DR EMBL; AE008809; AAL21329.1; -;  
DR StyGene; SG10475; pdxK.  
DR InterPro; IPR002173; pfkB.  
DR Pfam; PF00294; pfkB; 1.  
KW Transferase; Kinase; Zinc; Magnesium; Complete proteome.  
FT CONFLICT 241 241 A -> P (IN REF. 1).  
FT CONFLICT 283 283 MISSING (IN REF. 1).  
SQ SEQUENCE 288 AA; 30967 MW; CF3F701FEA841F7A CRC64;

Query Match 1.0%; Score 8; DB 1; Length 288;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 187 VTAVPTVL 194  
|||||||  
DB 51 VTAVPTVL 58

## RESULT 21

TLR6\_HUMAN STANDARD; PRT; 796 AA.  
ID TLR6\_HUMAN  
AC Q9Y2C9;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Toll-like receptor 6 precursor.  
GN TLR6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=99250250; PubMed=10231569;  
RA Takeuchi O., Kawai T., Sanjo H., Copeland N.G., Gilbert D.J.,  
RA Jenkins N.A., Takeda K., Akira S.;  
RT "TLR6: A novel member of an expanding Toll-like receptor family.";  
RL Gene 231:59-65(1999).  
RN [2]  
RP FUNCTION.  
RX PubMed=11441107;  
RA Bulut Y., Faure E., Thomas L., Equils O., Arditi M.;  
RT "Cooperation of Toll-like receptor 2 and 6 for cellular activation by  
RT soluble tuberculosis factor and Borrelia burgdorferi outer surface  
RT protein A lipoprotein: role of Toll-interacting protein and IL-1  
RT receptor signaling molecules in Toll-like receptor 2 signaling.";  
RL J. Immunol. 167:987-994(2001).  
CC -!- FUNCTION: Participates in the innate immune response to Gram-  
CC positive bacteria and fungi. Acts via MyD88 and TRAF6, leading to  
CC NF-kappa-B activation, cytokine secretion and the inflammatory  
CC response. Recognizes mycoplasma macrophage-activating  
CC lipopeptide-2kd (MALP-2) soluble tuberculosis factor (STF),  
CC phenol-soluble modulin (PSM) and B.burgdorferi outer surface  
CC protein A lipoprotein (OspA-L) cooperatively with TLR2.  
CC -!- SUBUNIT: Binds TLR2 via their respective extracellular domains.  
CC Binds MyD88 via their respective TIR domains (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Plasma membrane and  
CC phagosomes (By similarity).  
CC -!- TISSUE SPECIFICITY: Detected in monocytes, CD11c+ immature  
CC dendritic cells, plasmacytoid pre-dendritic cells and dermal  
CC microvessel endothelial cells.  
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -!- SIMILARITY: CONTAINS 13 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AB020807; BAA78631.1; -  
DR MIM; 605403; -  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_term.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR000157; TIR.  
DR Pfam; PF00560; LRR; 8.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; LRR; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 2.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00255; TIR; 1.  
DR PROSITE; PS0104; TIR; 1.  
KW Receptor; Immune response; Inflammatory response; Signal;  
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.  
FT SIGNAL 1 31  
FT CHAIN 32 796  
FT DOMAIN 32 586  
FT TRANSMEM 587 607  
FT DOMAIN 608 796  
FT REPEAT 51 74  
FT REPEAT 75 98  
FT REPEAT 100 120  
FT REPEAT 121 144  
FT REPEAT 158 175  
FT REPEAT 176 201  
FT REPEAT 222 248  
FT REPEAT 376 400  
FT REPEAT 402 428  
FT REPEAT 430 447  
FT REPEAT 449 472  
FT REPEAT 474 496  
FT REPEAT 498 517  
FT DOMAIN 640 784  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 359 359 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 434 434 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 583 583 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 796 AA; 91889 MW; 35CEAEC05BFB88BD CRC64;

Query Match 1.0%; Score 8; DB 1; Length 796;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 675 LDLSHNL 682  
Db 105 LDLSHNL 112

RESULT 22

BIMA\_EMENI  
ID BIMA\_EMENI STANDARD; PRT; 806 AA.  
AC P17885;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein bima.  
CN BIMA.  
OS Emericella nidulans (Aspergillus nidulans).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; Emericella.  
OX NCBI\_TaxID=5072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI53;  
RX MEDLINE=921121243; PubMed=1770001;  
RA O'Donnell K.L., Osmani A.H., Osmani S.A., Morris N.R.;  
RT "bima encodes a member of the tetratricopeptide repeat family of  
RT proteins and is required for the completion of mitosis in Aspergillus  
RT nidulans.";  
RL J. Cell Sci. 99:711-719(1991).  
CC -1- FUNCTION: REQUIRED FOR THE COMPLETION OF MITOSIS IN ASPERGILLUS  
CC NIDULANS.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- SIMILARITY: BELONGS TO THE CDC27/NUC2 FAMILY.  
CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X59269; CAA41959.1; -  
DR PIR; S21860; S21860.  
DR PIR; A53256; A53256.  
DR InterPro; IPR001440; TPR.  
DR Pfam; PF00515; TPR; 7.  
DR SMART; SM00028; TPR; 7.  
KW Cell division; Cell cycle; Mitosis; Repeat; TPR repeat;  
KW Nuclear protein.  
FT REPEAT 76 109 TPR 1.  
FT REPEAT 127 160 TPR 2.  
FT DOMAIN 260 399 BIMA DOMAIN.  
FT REPEAT 513 546 TPR 3.  
FT REPEAT 581 614 TPR 4.  
FT REPEAT 616 648 TPR 5.  
FT REPEAT 649 682 TPR 6.  
FT REPEAT 684 716 TPR 7.  
FT REPEAT 751 784 TPR 8.  
SQ SEQUENCE 806 AA; 89714 MW; F137BDE3A74C0457 CRC64;  
  
Query Match 1.0%; Score 8; DB 1; Length 806;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 259 ALTELKVL 266  
Db 736 ALTELKVL 743

RESULT 23  
SVL\_BORBU  
ID SVL\_BORBU STANDARD; PRT; 1042 AA.  
AC 051773;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--tRNA ligase)  
DE (IleRS).  
GN ILES OR BB0833.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE=98065943; PubMed=9403685;  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,

RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
RA Uterback T., Watney L., McDonald L., Artiach P., Bowman C.,  
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
RA Smith H.O., Venter J.C.;  
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
RT burgdorferi.";  
RL Nature 390:580-586(1997).  
CC -1- CATALYTIC ACTIVITY: ATP + L-isoleucine + tRNA(Ile) = AMP +  
CC diphosphate + L-isoleucyl-tRNA(Ile).  
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE001181; AAC67179.1; -;  
DR HSP; P56690; IILE.  
DR TIGR; BB0833; -;  
DR InterPro; IPR002300; tRNA-synt\_1a.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002301; tRNA-synt\_1le.  
DR Pfam; PF00133; tRNA-synt\_1; 1.  
DR PRINTS; PR00984; TRNASYNTHILE.  
DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; FALSE\_NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Metal-binding; Zinc; Complete proteome.  
FT SITE 48 58 "HIGH" REGION.  
FT SITE 594 598 "KMSKS" REGION.  
FT BINDING 597 597 ATP (BY SIMILARITY).  
SQ SEQUENCE 1042 AA; 122331 MW; 6C0F7D820CA32F75 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 1042;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 649 KNLSLAKN 656  
|||||||  
Db 670 KNLSLAKN 677

RESULT 24  
CAT8\_YEAST  
ID CAT8\_YEAST STANDARD; PRT; 1433 AA.  
AC P39113;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Regulatory protein CAT8.  
GN CAT8 OR MSP8 OR YMR280C OR YMR8021.06C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95198709; PubMed=7891685;  
RA Hedges D., Proft M., Entian K.-D.;  
RT "CAT8, a new zinc cluster-encoding gene necessary for derepression of  
RT gluconeogenic enzymes in the yeast Saccharomyces cerevisiae.";  
RL Mol. Cell. Biol. 15:1915-1922(1995).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=S288C / AB972;

RA Pearson D., Bowman S., Barrell B.G., Rajandream M.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RC SEQUENCE FROM N.A.  
RP STRAIN=ENY.WA-1A;  
RC Boles E., Hettmann C., Zimmermann F.K.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: ACTIVATOR OF THE GLUCONEOGENIC ENZYMES FBP1 AND PCK1  
CC GENES.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- PTM: COULD BE THE TARGET OF THE SNF1/CAT1 - SNF4/CAT3 KINASE  
CC COMPLEX.  
CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
CC CLUSTER DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X78344; CAA55139.1; -;  
DR EMBL; Z49704; CAA89778.1; -;  
DR EMBL; X94215; CAA63906.1; -;  
DR PIR; S48234; S48234.  
DR HSP; P04386; IAW6.  
DR TRANSFAC; T03227; -;  
DR SGD; S0004893; CAT8.  
DR InterPro; IPR001138; ZN2\_Cy6\_fungal.  
DR Pfam; PF00172; ZN\_c1us; 1.  
DR PRINTS; PR00054; FUNGALZNCYS.  
DR SMART; SM00066; GAL4; 1.  
DR PROSITE; PS00463; ZN2\_Cy6\_FUNGAL\_1; 1.  
DR PROSITE; PS00048; ZN2\_Cy6\_FUNGAL\_2; 1.  
KW Transcription regulation; DNA-binding; Nuclear protein; Zinc;  
KW Metal-binding; Phosphorylation; Carbohydrate metabolism; Activator.  
FT DNA\_BIND 70 97 ZN(2)-CYS(6), FUNGAL-TYPE.  
FT DOMAIN 208 212 POLY-PRO.  
FT DOMAIN 972 976 POLY-ASN.  
FT CONFLICT 303 303 S -> A (IN REF. 3).  
FT CONFLICT 747 747 K -> L (IN REF. 1).  
FT CONFLICT 768 768 T -> A (IN REF. 3).  
FT CONFLICT 999 1002 MISSING (IN REF. 3).  
FT CONFLICT 1008 1008 G -> S (IN REF. 3).  
FT CONFLICT 1016 1016 H -> Q (IN REF. 3).  
FT CONFLICT 1019 1019 Q -> P (IN REF. 3).  
FT CONFLICT 1061 1061 V -> M (IN REF. 3).  
FT CONFLICT 1072 1072 T -> A (IN REF. 3).  
FT CONFLICT 1074 1074 A -> S (IN REF. 3).  
FT CONFLICT 1092 1092 I -> V (IN REF. 3).  
FT CONFLICT 1100 1100 I -> V (IN REF. 3).  
FT CONFLICT 1120 1120 M -> L (IN REF. 3).  
FT CONFLICT 1162 1162 S -> A (IN REF. 3).  
SQ SEQUENCE 1433 AA; 160485 MW; 5ED790BEFB47B632 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 1433;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 PSLQLSL 126  
|||||||  
Db 520 PSLQLSL 527

RESULT 25  
RL36\_RICPR  
ID RL36\_RICPR STANDARD; PRT; 41 AA.  
AC Q9ZD87;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)

```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L36.
GN RPMJ OR RP456.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
CC -1- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ235271; CAAL4912.1; -
DR InterPro: IPR000473; Ribosomal_L36.
DR Pfam: PF00444; Ribosomal_L36; 1.
DR PROSITE: PS00828; RIBOSOMAL_L36; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 41 AA; 4861 MW; EEBB71BC7C7A5C76 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 SSLSLK 344
DB 5 SSLSLK 11

RESULT 26
NULM_BRALA
ID NULM_BRALA STANDARD; PRT; 91 AA.
AC Q79420; Q47424;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN ND4L OR NAD4L OR NADH4L.
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus), and
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7740, 7739;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=B. lanceolatum;
RX MEDLINE=98292550; PubMed=9628930;
RA Spruyt N., Delarbre C., Gachelin G., Laudet V.;
RT "Complete sequence of the amphioxus (Branchiostoma lanceolatum)
RT mitochondrial genome: relations to vertebrates.";
RL Nucleic Acids Res. 26:3279-3285(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=B. floridae;
RX MEDLINE=99261652; PubMed=10331267;
RA Boore J.L., Daehler L.L., Brown W.M.;
RT "Complete sequence, gene arrangement, and genetic code of

16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L36.
GN RPMJ OR RP456.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
CC -1- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ235271; CAAL4912.1; -
DR InterPro: IPR000473; Ribosomal_L36.
DR Pfam: PF00444; Ribosomal_L36; 1.
DR PROSITE: PS00828; RIBOSOMAL_L36; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 41 AA; 4861 MW; EEBB71BC7C7A5C76 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 SSLSLK 344
DB 5 SSLSLK 11

RESULT 26
NULM_BRALA
ID NULM_BRALA STANDARD; PRT; 91 AA.
AC Q79420; Q47424;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN ND4L OR NAD4L OR NADH4L.
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus), and
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7740, 7739;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=B. lanceolatum;
RX MEDLINE=98292550; PubMed=9628930;
RA Spruyt N., Delarbre C., Gachelin G., Laudet V.;
RT "Complete sequence of the amphioxus (Branchiostoma lanceolatum)
RT mitochondrial genome: relations to vertebrates.";
RL Nucleic Acids Res. 26:3279-3285(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=B. floridae;
RX MEDLINE=99261652; PubMed=10331267;
RA Boore J.L., Daehler L.L., Brown W.M.;
RT "Complete sequence, gene arrangement, and genetic code of

mitochondrial DNA of the cephalochordate Branchiostoma floridae
(RAmphioxus).";
RL Mol. Biol. Evol. 16:410-418(1999).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y16474; CAA76255.1; -
DR EMBL: AF098298; AAB87993.2; -
DR InterPro: IPR003214; Mit_NADH_oxidoreductse_4L.
DR InterPro: IPR001133; Oxidored_q2.
DR Pfam: PF00420; oxidored_q2; 1.
DR ProDom: PD000359; Mit_NADH_oxidoreductse_4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 91 AA; 9751 MW; 7E7D093F02468BD3 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 356 LKSFNLS 362
DB 85 LKSFNLS 91

RESULT 27
ARRL_ECOLI
ID ARRL_ECOLI STANDARD; PRT; 117 AA.
AC P15905;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Arsenical resistance operon repressor.
GN ARSR.
OS Escherichia coli.
OG Plasmid R773.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174986; PubMed=2408017;
RA San Francisco M.J.D., Hope C.L., Owolabi J.B., Tisa L.S., Rosen B.P.;
RT "Identification of the metalloregulatory element of the
RT plasmid-encoded arsenical resistance operon.";
RL Nucleic Acids Res. 18:619-624(1990).
RN [2]
RP FUNCTION.
RX MEDLINE=92157859; PubMed=1838573;
RA Wu J., Rosen B.P.;
RT "The Arsr protein is a trans-acting regulatory protein.";
RL Mol. Microbiol. 5:1331-1336(1991).
RN [3]
RP METAL-REGULATION.
RX MEDLINE=93107054; PubMed=8416957;
RA Wu J., Rosen B.P.;
RT "Metalloregulated expression of the ars operon.";
RL J. Biol. Chem. 268:52-58(1993).
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE ARS OPERON. ARSR IS
CC A TRANS-ACTING REGULATORY PROTEIN WHICH CONTROLS ITS OWN
CC EXPRESSION. THE REPRESSIVE EFFECT OF ARSR IS ALLEVIATED BY OXYIONS
CC OF +III OXIDATION STATE OF ARSENIC, ANTIMONY, AND BISMUTH, AS WELL
CC AS ARSENATE (AS(V)).
CC -1- SUBUNIT: BINDS DNA AS AN HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
```

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL; X16045; CAA341168.1; -.  
 DR PIR; JS0448; BVECAR.  
 DR HSP; P30340; 1SMT.  
 DR InterPro; IPR001845; HTH\_ARSR.  
 DR Pfam; PF01022; HTH\_5; 1.  
 DR PRINTS; PR00778; HTHARSR.  
 DR SMART; SM00418; HTH\_ARSR; 1.  
 KW Plasmid; Arsenical resistance; Transcription regulation; Repressor;  
 KW DNA-binding. 33 52 H-T-H MOTIF (POTENTIAL).  
 FT DNA\_BIND 117 AA; 13198 MW; 1F0D10766E4FD886 CRC64;  
 SQ SEQUENCE 117 AA; 13198 MW; 1F0D10766E4FD886 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 614 LQLFKNL 620  
 |||||  
 DB 7 LQLFKNL 13

RESULT 28  
 YG3E\_YEAST  
 ID YG3E\_YEAST STANDARD; PRT; 124 AA.  
 AC P53282;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Hypothetical 14.5 kDa protein in PAS5-CBF2 intergenic region.  
 GN YGR137W  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Talla E.,  
 RA Nawrocki A., del Bino S., Goffeau A.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL; Z72921; CAA97150.1; -.  
 DR SGD; S0003369; YGR137W.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 70 90 POTENTIAL.  
 SQ SEQUENCE 124 AA; 14462 MW; BC11E613AFD1E189 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 266 LRLHSNS 272  
 |||||  
 DB 10 LRLHSNS 16

RESULT 29  
 VG61\_BPML5  
 ID VG61\_BPML5 STANDARD; PRT; 125 AA.  
 AC Q05274;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE Gene 61 protein (GP61).  
 GN 61.  
 OS Mycobacteriophage L5.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=31757;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93211282; PubMed=8459766;  
 RA Hatfull G.F., Sarkis G.J.;  
 RT "DNA sequence, structure and gene expression of mycobacteriophage L5:  
 RT a phage system for mycobacterial genetics.";  
 RL Mol. Microbiol. 7:395-405(1993).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL; Z18946; CAA79437.1; -.  
 DR PIR; S31006; S31006.  
 -SQ SEQUENCE 125 AA; 14176 MW; 541035C6EE6647FB CRC64;

Query Match 0.9%; Score 7; DB 1; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 RSFSGLT 98  
 |||||  
 DB 28 RSFSGLT 34

RESULT 30  
 R1MM\_RICPR  
 ID R1MM\_RICPR STANDARD; PRT; 165 AA.  
 AC Q92D10;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable 16S rRNA processing protein r1mm.  
 GN R1MM OR R2348.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MDRID E;  
 RX MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria.";  
 RL Nature 396:133-140(1998).  
 CC -!- FUNCTION: ESSENTIAL FOR EFFICIENT PROCESSING OF 16S rRNA. PROBABLY  
 CC PART OF THE 30S SUBUNIT PRIOR TO OR DURING THE FINAL STEP IN THE  
 CC PROCESSING OF 16S rRNA. 30S RIBOSOMAL SUBUNIT. IT COULD BE SOME  
 CC ACCESSORY PROTEIN NEEDED FOR EFFICIENT ASSEMBLY OF THE 30S  
 CC SUBUNIT. R1MM IS NEEDED IN A STEP PRIOR TO RBFA DURING THE  
 CC MATURATION OF 16S rRNA. HAS AFFINITY FOR FREE RIBOSOMAL 30S

```

CC SUBUNITS BUT NOT FOR 70S RIBOSOMES (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -I- SIMILARITY: BELONGS TO THE R1MM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ235271; CAA14808.1; -.
DR InterPro; IPR002676; R1mm.
DR Pfam; PF01782; R1mm; 1.
KW rRNA processing; Complete proteome.
SQ SEQUENCE 165 AA; 18772 MW; 25DDB377BC250919 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 695 SLKNLIL 701
DB 3 SLKNLIL 9

RESULT 31
RELX_MESAU
ID RELX_MESAU STANDARD; PRT; 177 AA.
AC Q64171;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prorelaxin precursor.
GN R1N.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96115021; PubMed=7492700;
RA McCaslin R.B., Renegar R.H.;
RT "Determination of the prorelaxin nucleotide sequence and expression
RL of prorelaxin messenger ribonucleic acid in the golden hamster.";
RL Biol. Reprod. 53:454-461(1995).
CC -I- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
CC PRODUCE DILATATION OF THE BIRTH CANAL IN MANY MAMMALS. IT BEARS
CC MATURE YOUNG, AND ALLOWS SEPARATION OF THE PELVIC BONES.
CC -I- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S79879; AAB35655.1; -.
DR InterPro; IPR000739; Insulin_IGF_relaxin.
DR Pfam; PF00049; Insulin; 1.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Signal.
FT SIGNAL 1 22 BY SIMILARITY.

```

```

FT CHAIN 23 59 RELAXIN B CHAIN (PROBABLE).
FT PROPEP 64 149 CONNECTING PEPTIDE (PROBABLE).
FT CHAIN 154 177 RELAXIN A CHAIN (PROBABLE).
FT DISULFID 36 164 INTERCHAIN (BY SIMILARITY).
FT DISULFID 48 177 INTERCHAIN (BY SIMILARITY).
FT DISULFID 163 168 BY SIMILARITY.
SQ SEQUENCE 177 AA; 20007 MW; 6925562BD8C66CCD CRC64;

Query Match 0.9%; Score 7; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LKSLYLD 106
DB 140 LKSLYLD 146

RESULT 32
YC47_METJA
ID YC47_METJA STANDARD; PRT; 180 AA.
AC Q58644;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1247.
GN MJ1247.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sutton G.G., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Kierulff A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC -I- SIMILARITY: BELONGS TO THE SIS FAMILY. YCKF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67565; AAB99251.1; -.
DR TIGR; MJ1247.
DR InterPro; IPR001347; SIS.
DR Pfam; PF01380; SIS; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 180 AA; 20443 MW; 7C3D607BCBD4AA0A CRC64;

Query Match 0.9%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 622 KLEELDI 628
DB 3 KLEELDI 9

```

```
RESULT 33
Y628_CHLMU STANDARD; PRT; 196 AA.
AC Q9PK45;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Maf-like protein TC0628.
GN TC0628.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Ni9g.
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bowman C.,
RA Linher K., Weidman J., Khouri H., Craven B., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."; 28:1397-1406(2000).
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE MAF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE022331; AAF39457.1; -.
DR TIGR; TC0628; -.
DR InterPro; IPR003697; Maf.
DR Pfam; PF02545; Maf; 1.
KW Complete proteome.
FT ACT_SITE 34
FT ACT_SITE 34
SQ SEQUENCE 196 AA; 21730 MW; 9886B059E52D78A5 CRC64;
      1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
      31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
      60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
      89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113
      114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
      136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157
      158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
      180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196

Query Match 0.9%; Score 7; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 QAFSSLK 341
DB 141 QAFSSLK 147

RESULT 34
TSAA_BUCAI STANDARD; PRT; 197 AA.
AC P57279;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable peroxiredoxin.
GN BU182.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOKYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
```

```
RL Nature 407:81-86(2000).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001118; BAB12899.1; -.
DR InterPro; IPR000866; Ahpc-TSA.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Antioxidant; Complete proteome.
FT ACT_SITE 50
FT ACT_SITE 50
SQ SEQUENCE 197 AA; 22363 MW; D08C2CA66C256376 CRC64;
      1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
      31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
      60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
      89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113
      114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
      136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157
      158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
      180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197

Query Match 0.9%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 700 ILKNNQI 706
DB 16 ILKNNQI 22

RESULT 35
VP24_NPVAC STANDARD; PRT; 198 AA.
ID VP24_NPVAC
AC P41678;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Capsid protein P24.
GN P24.
OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus.";
RL Virology 202:586-605(1994).
CC -!- SIMILARITY: TO THE CORRESPONDING PROTEIN OF OPMNPV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L22858; AAA66759.1; -.
DR Coar protein; Late protein.
KW Coar protein; Late protein.
SQ SEQUENCE 198 AA; 22110 MW; 4019AC5E8442D7D2 CRC64;
      1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
      31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
      60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88
      89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113
      114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
      136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157
      158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
      180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198

Query Match 0.9%; Score 7; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 INKLOEL 290
DB 158 INKLOEL 164
```

```
RESULT 36
ERP5_YEAST STANDARD; PRT; 212 AA.
ID P38819.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ERP5 protein precursor.
GN ERP5 OR YHR110W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevasilis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
RL Science 265:2077-2082(1994).
CC -!- FUNCTION: INVOLVED IN VESICULAR PROTEIN TRAFFICKING (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC reticulum (By similarity).
CC -!- SIMILARITY: BELONGS TO THE EMP24/GP25L FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U00059; AAB68853.1; -.
DR PIR: S48952; S48952.
DR SGD: S0001152; ERP5.
DR InterPro: IPR000348; Emp24_gp25L_p24.
DR Pfam: PF01105; EMP24_GP25L; 1.
KW Transport; Protein transport; Transmembrane; Signal;
KW Endoplasmic reticulum; Glycoprotein.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 212 ERP5 PROTEIN.
FT DOMAIN 21 178 LUMENAL (POTENTIAL).
FT TRANSMEM 179 199 POTENTIAL.
FT DOMAIN 200 212 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 171 171 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 212 AA; 24248 MW; 5602610D31AE04D1 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 212;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 717 AFQLRYL 723
Db 196 AFQLRYL 202
|||||

RESULT 37
RPOA_RECAM STANDARD; PRT; 222 AA.
ID RPOA_RECAM
AC O21260.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (Transcriptase
DE alpha chain) (RNA polymerase alpha subunit).
GN RPOA.
OS Reclinomonas americana.
OG Mitochondrion.
OC Eukaryota; core jakobids; Reclinomonas.
OX NCBI_TaxID=48483;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 50394; PubMed=9168110;
RX MEDLINE=97311393; PubMed=9168110;
RA Lang B.F., Burger G., O'Kelly C.J., Cedergren R., Golding G.B.,
RA Lemieux C., Sankoff D., Turmel M., Gray M.W.;
RT "An ancestral mitochondrial DNA resembling a eubacterial genome in
RT miniature."
RL Nature 387:493-497(1997).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE ALPHA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF007261; AAD11887.1; -.
DR InterPro: IPR001700; RNA_pol_A_bac.
DR Pfam: PF01000; RNA_pol_A_bac; 1.
DR ProDom: PD001179; RNA_pol_A_bac; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription;
KW Mitochondrion.
SQ SEQUENCE 222 AA; 25375 MW; 1D48F788696CB79A CRC64;

Query Match 0.9%; Score 7; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 741 NVLNNLK 747
Db 80 NVLNNLK 86
|||||

RESULT 38
DEOC_MYCPN STANDARD; PRT; 224 AA.
ID DEOC_MYCPN
AC P09924.
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyriboaldolase)
DE (Deoxyriboaldolase).
GN DEOC OR MPN063 OR MP091.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=89128453; PubMed=2492658;
RA Loechel S., Inamine J.M., Hu P.-C.;
RT "Nucleotide sequence of the deoc gene of Mycoplasma pneumoniae."
RL Nucleic Acids Res. 17:801-801(1989).
RN [2]
RP SEQUENCE FROM N.A.
```



RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
pneumoniae";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate = D-  
glyceraldehyde 3-phosphate + acetaldehyde.  
CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.  
CC DEOC SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X13544; CAA31897.1; -;  
DR EMBL; AE000011; AAB95739.1; -;  
DR PIR; S02216; S02216.  
DR InterPro; IPR002915; Deoc.  
DR Pfam; PF01791; Deoc; 1.  
KW Lyase; Schiff base; Complete proteome.  
FT BINDING 152 152  
SQ SEQUENCE 224 AA; 24878 MW; 73C3E4932E7881F7 CRC64;  
  
Query Match 0.9%; Score 7; DB 1; Length 224;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 534 EELHLE 540  
Db 213 EELHLE 219  
|||||||  
  
RESULT 39  
YMI7\_YEAST STANDARD; PRT; 237 AA.  
AC P40207;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Hypothetical 27.9 kDa protein in REC114-PS02 intergenic region.  
GN YMR134W OR YW9375.03.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 247071; CAA87348.1; -;  
DR SGD; S0004741; YMR134W.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 36 56  
SQ SEQUENCE 237 AA; 27921 MW; 9DCF3CED15B4A622 CRC64;  
  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
pneumoniae";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate = D-  
glyceraldehyde 3-phosphate + acetaldehyde.  
CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.  
CC DEOC SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X13544; CAA31897.1; -;  
DR EMBL; AE000011; AAB95739.1; -;  
DR PIR; S02216; S02216.  
DR InterPro; IPR002915; Deoc.  
DR Pfam; PF01791; Deoc; 1.  
KW Lyase; Schiff base; Complete proteome.  
FT BINDING 152 152  
SQ SEQUENCE 224 AA; 24878 MW; 73C3E4932E7881F7 CRC64;  
  
Query Match 0.9%; Score 7; DB 1; Length 224;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 534 EELHLE 540  
Db 213 EELHLE 219  
|||||||  
  
RESULT 39  
YMI7\_YEAST STANDARD; PRT; 237 AA.  
AC P40207;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Hypothetical 27.9 kDa protein in REC114-PS02 intergenic region.  
GN YMR134W OR YW9375.03.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 247071; CAA87348.1; -;  
DR SGD; S0004741; YMR134W.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 36 56  
SQ SEQUENCE 237 AA; 27921 MW; 9DCF3CED15B4A622 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 237;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 284 INKLQEL 290  
Db 15 INKLQEL 21  
|||||||  
  
RESULT 40  
CB21\_PINTH STANDARD; PRT; 266 AA.  
ID CB21\_PINTH  
AC P10049;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Chlorophyll A-B binding protein type I, chloroplast precursor (CAB)  
(LHCP).  
OS Pinus thunbergii (Green pine) (Japanese black pine).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_TaxID=3350;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89098344; PubMed=3211759;  
RA Yanamoto N., Matsuoka M., Kano Murakami Y., Tanaka Y., Ohashi Y.;  
RT "Nucleotide sequence of a full length cDNA clone of ribulose  
biphosphate carboxylase small subunit gene from green dark-grown  
pine (Pinus thunbergii) seedling";  
RL Nucleic Acids Res. 16:11830-11830(1988).  
CC -1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT  
RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS  
WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN  
EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF  
GRANAL MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION  
OF ITS THREONINE RESIDUES; BOTH ARE BELIEVED TO MEDIATE THE  
DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.  
CC -1- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND  
CHLOROPHYLL A-B BINDING PROTEINS.  
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X13407; CAA31773.1; -;  
DR PIR; S02045; S02045.  
DR InterPro; IPR001344; Chloro\_AB\_bind.  
DR Pfam; PF00504; chloroa\_b\_bind; 1.  
KW Chlorophyll; Photosynthesis; Photosystem I; Photosystem II;  
KW Thylakoid; Membrane; Chloroplast; Transit peptide; Multigene family;  
KW Transmembrane; Phosphorylation.  
FT TRANSIT 1 37  
FT CHAIN 38 266  
FT TRANSMEM 100 119  
FT TRANSMEM 152 172  
FT TRANSMEM 220 236  
SQ SEQUENCE 266 AA; 28543 MW; DFB592FC60420659 CRC64;  
  
Query Match 0.9%; Score 7; DB 1; Length 266;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 653 LAKNGLK 659  
Db 119 LAKNGLK 125  
|||||||



```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AE000244; AAD13435.1; -
DR EcoGene: EG13782; yddK.
DR InterPro: IPR001611; LRR.
DR Pfam: PF00560; LRR; 5.
KW Hypothetical protein; Repeat; Leucine-rich repeat; Complete proteome.
FT REPEAT 10 31
FT REPEAT 32 52
FT REPEAT 107 130
FT REPEAT 132 151
FT REPEAT 152 174
FT REPEAT 193 214
FT REPEAT 215 237
FT REPEAT 258 281
FT REPEAT 282 305
FT REPEAT LRR 1.
FT REPEAT LRR 2.
FT REPEAT LRR 3.
FT REPEAT LRR 4.
FT REPEAT LRR 5.
FT REPEAT LRR 6.
FT REPEAT LRR 7.
FT REPEAT LRR 8.
FT REPEAT LRR 9.
SQ SEQUENCE 318 AA; 36241 MW; B2ED1E202A43A753 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 677 LSHNQLT 683
Db 201 LSHNQLT 207

RESULT 44
ERMS_STRFR STANDARD; PRT; 319 AA.
AC P45439;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE RNA adenine N-6-methyltransferase (EC 2.1.1.48) (Macrolide-
DE lincosamide-streptogramin B resistance protein) (Brythromycin
DE resistance protein).
GN ERMSF OR TLRA.
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=88169508; PubMed=3127381;
RA Kamimura S., Weisblum B.;
RT Translational attenuation control of ermsF, an inducible resistance
RT determinant encoding rRNA N-methyltransferase from Streptomyces
RT fradiae.;
RL J. Bacteriol. 170:1800-1811(1988).
CC -!- FUNCTION: THIS PROTEIN PRODUCES A DIMETHYLATION OF THE ADENINE
CC RESIDUE AT POSITION 2058 IN 23S rRNA, RESULTING IN REDUCED
CC AFFINITY BETWEEN RIBOSOMES & MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B
CC ANTIBIOTICS.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing N6-methyladenine.
CC -!- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
-----
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M19269; AAA26742.1; -
DR InterPro: IPR001737; RNA_A_dimeth.
DR Pfam: PF00398; RnaAD; 1.
DR PROSITE: PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Transferase; Methyltransferase.
SQ SEQUENCE 319 AA; 35527 MW; 3A543FA222CFB7DB CRC64;

Query Match 0.9%; Score 7; DB 1; Length 319;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 292 LSONFLA 298
Db 63 LSONFLA 69

RESULT 45
YECF_ECOLI STANDARD; PRT; 323 AA.
ID YECF_ECOLI
AC P76291; O07983;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yecP.
GN YECF OR Bi871.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasubramanian S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
CC -!- SIMILARITY: STRONG, TO H.INFLUENZAE H11351.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AE000280; AAC74941.1; -
DR EMBL: D90829; BAA15681.1; -
DR EcoGene: EG14034; yecP.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 323 AA; 37006 MW; 2954B076A83607A3 CRC64;
```

Query Match 0.9%; Score 7; DB 1; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 524 RDLHLS 530  
 |||||  
 Db 56 RDLHLS 62 \*

## RESULT 46

MDH\_THEFL  
 ID MDH\_THEFL STANDARD; PRT; 327 AA.  
 AC P10584;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Malate dehydrogenase (EC 1.1.1.37).  
 GN MDH.  
 OS Thermus aquaticus (subsp. flavus).  
 OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.  
 OX NCBI\_TaxID=274;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-37 AND 265-284.  
 RC STRAIN=ATCC 33923 / AT-62;  
 RX MEDLINE=87033605; PubMed=3771528;  
 RA Nishiyama M., Matsubara N., Yamamoto K., Iijima S., Uozumi T.,  
 RA Beppu T.;  
 RT "Nucleotide sequence of the malate dehydrogenase gene of Thermus  
 RT flavus and its mutation directing an increase in enzyme activity.";  
 RL J. Biol. Chem. 261:14178-14183(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33923 / AT-62;  
 RX MEDLINE=91238680; PubMed=2034208;  
 RA Nishiyama M., Horinouchi S., Beppu T.;  
 RT "Characterization of an operon encoding succinyl-CoA synthetase and  
 RT malate dehydrogenase from Thermus flavus AT-62 and its expression in  
 RT Escherichia coli.";  
 RL Mol. Gen. Genet. 226:1-9(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B / NCBI 11247;  
 RX MEDLINE=90375010; PubMed=2204576;  
 RA Nicholls D.J., Sundaram T.K., Atkinson T., Minton N.P.;  
 RT "Cloning and nucleotide sequences of the mdh and sucD genes from  
 RT Thermus aquaticus B.";  
 RL FEBS Microbiol. Lett. 58:7-14(1990).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RC STRAIN=ATCC 33923 / AT-62;  
 RX MEDLINE=93229488; PubMed=8471603;  
 RA Kelly C.A., Nishiyama M., Ohnishi Y., Beppu T., Birktoft J.J.;  
 RT "Determinants of protein thermostability observed in the 1.9-A  
 RT bacterial structure of malate dehydrogenase from the thermophilic  
 RT bacterium Thermus flavus.";  
 RL Biochemistry 32:3913-3922(1993).  
 CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 CC EMBL: J02598; AAA27499.1; -;  
 CC EMBL: X54073; CAA38008.1; -;  
 CC EMBL: X56033; CAA39508.1; -;  
 CC EMBL: A26565; CAA01826.1; -;

DR PIR: A26065; DETWMA.  
 DR PIR: S12139; DETWMB.  
 DR PDB: 1BMD; 31-JUL-94.  
 DR PDB: 1BDM; 20-DEC-94.  
 DR InterPro: IPR001252; MDH\_actsite.  
 DR InterPro: IPR001236; ldh.  
 DR Pfam: PF00056; ldh; 1.  
 DR Pfam: PF02866; ldh\_C; 1.  
 DR ProDom: PD003052; MDH\_actsite; 1.  
 DR PROSITE: PS00068; MDH; 1.  
 KW Oxidoreductase; Tricarboxylic acid cycle; NAD; 3D-structure.  
 FT ACT\_SITE 159 PROTON-RELAY.  
 FT BINDING 162 SUBSTRATE CARBOXYL GROUP.  
 FT ACT\_SITE 187 PROTON-RELAY.  
 FT VARIANT 190 T -> I (IN MUTANT STRAIN F428, PRODUCES  
 FT SEQUENCE 327 AA; 35426 MW; 31FA90DED2393DF2 CRC64;  
 SQ

Query Match 0.9%; Score 7; DB 1; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 QLLEIPQ 115  
 |||||  
 Db 39 QLLEIPQ 45

## RESULT 47

PARB\_MYCLE  
 ID PARB\_MYCLE STANDARD; PRT; 333 AA.  
 AC Q50201;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable chromosome partitioning protein parB.  
 GN PARB OR ML2706.  
 OS Mycobacterium leprae.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97124199; PubMed=8969512;  
 RA Fsihi H., De Rossi E., Salazar L., Cantoni R., Labo M., Riccardi G.,  
 RA Takiff H.E., Eiglmeier K., Bergh S., Cole S.T.;  
 RT "Gene arrangement and organization in a approximately 76 kb fragment  
 RT encompassing the oric region of the chromosome of Mycobacterium  
 RT leprae.";  
 RL Microbiology 142:3147-3161(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus.";  
 RL Nature 409:1007-1011(2001).  
 CC -1- FUNCTION: INVOLVED IN CHROMOSOME PARTITION. LOCALIZE TO BOTH POLES  
 CC OF THE PREDIVISIONAL CELL FOLLOWING COMPLETION OF DNA REPLICATION.  
 CC BINDS TO THE DNA ORIGIN OF REPLICATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE PARB FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; L39923; AAB53134.1; -  
DR EMBL; AL583926; CAC32238.1; ALT\_INIT.  
DR Leproma; ML2706; -  
DR InterPro; IPR003115; ParBc.  
DR Pfam; PF02195; ParBc; 1.  
DR SMART; SM00470; ParBc; 1.  
KW Chromosome partition; DNA-binding; Complete proteome.  
SQ SEQUENCE 333 AA; 36350 MW; 9AF78955A73769A2 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 333;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 LLSLEAN 129  
| | | | |  
Db 224 LLSLEAN 230

RESULT 48  
RTCA\_METJA STANDARD; PRT; 338 AA.  
ID RTCA\_METJA  
AC O60335;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (RNA-3'-  
DE phosphate cyclase) (RNA cyclase).  
GN RTCA OR MJ0025.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT \*Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.\*;  
RL Science 273:1058-1073(1996).

CC -1- FUNCTION: CATALYZES THE CONVERSION OF 3'-PHOSPHATE TO A 2',3'-  
CC CYCLIC PHOSPHODIESTER AT THE END OF RNA. THE MECHANISM OF ACTION  
CC OF THE ENZYME OCCURS IN 3 STEPS: (A) ADENYLATION OF THE ENZYME BY  
CC ATP; (B) THE ENZYME ACTS ON RNA-N3'P TO PRODUCE RNA-N3'PP5'A; (C)  
CC A NON CATALYTIC NUCLEOPHILIC ATTACK BY THE ADJACENT 2'-HYDROXYL ON  
CC THE PHOSPHORUS IN THE DIESTER LINKAGE TO PRODUCE THE CYCLIC END  
CC PRODUCT. THE BIOLOGICAL ROLE OF THIS ENZYME IS UNKNOWN BUT IT IS  
CC LIKELY TO FUNCTION IN SOME ASPECTS OF CELLULAR RNA PROCESSING (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate = AMP +  
CC diphosphate + RNA terminal-2',3'-cyclic-phosphate.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- SIMILARITY: BELONGS TO THE RNA 3'-TERMINAL CYCLASE FAMILY.  
CC SUBFAMILY 1.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; U67461; AAB98006.1; -  
DR HSSP; P46849; IQMI.  
DR TIGR; MJ0025; -  
DR InterPro; IPR000228; RTC.  
DR Pfam; PF01137; RTC; 1.  
DR PROSITE; PS01287; RTC; 1.  
KW Ligase; Complete proteome.  
FT ACT\_SITE 304 304 BY SIMILARITY.  
SQ SEQUENCE 338 AA; 37228 MW; 4681A2AB120FD6F0 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 338;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 260 LTELKVL 266  
| | | | |  
Db 148 LTELKVL 154

RESULT 49  
KITH\_VZVD  
ID KITH\_VZVD STANDARD; PRT; 341 AA.  
AC P09250;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Thymidine kinase (EC 2.7.1.21).  
GN K6.  
OS Varicella-zoster virus (strain Dumas) (VZV).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10338;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86306657; PubMed=3018124;  
RA Davison A.J., Scott J.E.;  
RT \*The complete DNA sequence of varicella-zoster virus.\*;  
RL J. Gen. Virol. 67:1759-1816(1986).  
CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-  
CC phosphate.  
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES THYMIDINE KINASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; X04370; CAA27919.1; -  
DR PIR; A27341; KIB3C.  
DR HSSP; P03176; 2K15.  
DR InterPro; IPR001889; TK\_herpes.  
DR Pfam; PF00693; TK\_herpes; 1.  
DR ProDom; PD001519; TK\_herpes; 1.  
KW Transferase; Kinase; DNA synthesis; ATP-binding.  
FT NP\_BIND 19 26 ATP (POTENTIAL).  
SQ SEQUENCE 341 AA; 37817 MW; 1C95CC39750B0C07 CRC64;

Query Match 0.9%; Score 7; DB 1; Length 341;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 689 LSNCSRS 695  
| | | | |  
Db 284 LSNCSRS 290

```
RESULT 50
T4BB_BACCO
ID T4BB_BACCO STANDARD; PRT; 341 AA.
AC Q07606;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Restriction enzyme BcgI beta subunit (EC 3.1.21.-) (S.BcgI).
GN BCGIB.
OS Bacillus coagulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1398;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 55055;
RX MEDLINE=93197166; PubMed=8451198;
RA Kong H., Morgan R.D., Maunus R.E., Schildkraut I.;
RT "A unique restriction endonuclease, BcgI, from Bacillus coagulans.";
RL Nucleic Acids Res. 21:987-991(1993).
RN [2]
RP SEQUENCE OF 1-18, AND CHARACTERIZATION.
RC STRAIN=ATCC 55055;
RX MEDLINE=94103292; PubMed=8276869;
RA Kong H., Roemer S.E., Waite-Rees P.A., Benner J.S., Wilson G.G.,
RA Nwankwo D.O.;
RT "Characterization of BcgI, a new kind of restriction-modification
RT system.";
RL J. Biol. Chem. 269:683-690(1994).
CC -1- FUNCTION: RECOGNIZES THE DOUBLE STRANDED SEQUENCE 5'-CGAN(6)TGC-3'
CC AND CLEAVES BILATERALLY AND SYMMETRICALLY OUTSIDE THE SEQUENCE TO
CC RELEASE A 34-BASE PAIR FRAGMENT. METHYLATION OF THE RECOGNITION
CC SEQUENCE OCCURS ON THE ADENINE IN EITHER ONE OR BOTH STRANDS.
CC -1- COFACTOR: MAGNESIUM.
CC -1- SUBUNIT: HETEROTRIMER OF TWO ALPHA AND ONE BETA SUBUNIT. BOTH
CC SUBUNITS ARE NECESSARY FOR DNA-BINDING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L17341; AAA16627.1; -.
DR REBASE; 4024; S.BcgI.
KW Hydrolase; Endonuclease; Nuclease; Restriction system; Magnesium.
SQ SEQUENCE 341 AA; 39161 MW; E803D59E548AFD40 CRC64;
```

Query Match 0.98; Score 7; DB 1; Length 341;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 340 LKSLKIL 346  
| | | | |  
Db 128 LKSLKIL 134

Search completed: July 17, 2002, 09:49:08  
Job time: 255 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2002, 09:44:08 ; Search time 14.85 seconds  
(without alignments)  
1327.371 Million cell updates/sec

Title: US-09-202-054-1\_COPY\_30\_836

Perfect score: 807

Sequence: 1 FPKTLPDVTLDVPRNHVIV.....PGAHRGQSVISLDLYTCELD 807

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 6

Total number of hits satisfying chosen parameters: 1029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.1	236	1	US-08-442-063A-42
2	9	1.1	282	1	US-08-442-063A-45
3	9	1.1	307	1	US-08-442-063A-48
4	9	1.1	320	1	US-07-613-083B-1
5	9	1.1	333	1	US-08-442-063A-27
6	9	1.1	342	1	US-08-272-919-2
7	9	1.1	342	1	US-08-619-916-2
8	9	1.1	342	5	PCT-US95-08542-2
9	9	1.1	359	1	US-08-303-238-4
10	9	1.1	359	4	US-08-458-834-4
11	9	1.1	610	1	US-07-821-717B-6
12	9	1.1	610	1	US-08-119-262B-6
13	9	1.1	610	1	US-08-135-929A-11
14	9	1.1	610	1	US-08-234-265A-11
15	8	1.0	65	4	US-08-353-585-8
16	8	1.0	268	4	US-09-353-585-6
17	8	1.0	283	4	US-09-171-461-23
18	8	1.0	806	4	US-08-945-983-2
19	8	1.0	863	2	US-08-666-271-2
20	8	1.0	1012	2	US-08-475-891A-4
21	8	1.0	1023	2	US-08-475-891A-2
22	8	1.0	1023	2	US-08-567-375-2
23	8	1.0	1023	2	US-08-587-680A-2
24	8	1.0	1025	2	US-08-567-375-4
25	8	1.0	1025	2	US-08-587-680A-4
26	7	0.9	99	2	US-08-710-749-13
27	7	0.9	99	2	US-08-710-749-14

28	7	0.9	99	2	US-08-710-749-16	Sequence 16, Appl
29	7	0.9	102	2	US-08-710-749-8	Sequence 8, Appl
30	7	0.9	154	4	US-09-228-986-99	Sequence 99, Appl
31	7	0.9	157	4	US-09-228-986-90	Sequence 90, Appl
32	7	0.9	160	4	US-09-117-257-38	Sequence 38, Appl
33	7	0.9	160	4	US-09-489-352-38	Sequence 38, Appl
34	7	0.9	172	2	US-08-923-738-2	Sequence 2, Appl
35	7	0.9	172	2	US-08-923-738-4	Sequence 4, Appl
36	7	0.9	172	4	US-08-936-165A-385	Sequence 385, App
37	7	0.9	192	4	US-09-475-316A-23	Sequence 23, Appl
38	7	0.9	194	4	US-09-117-257-17	Sequence 17, Appl
39	7	0.9	194	4	US-08-945-476-17	Sequence 17, Appl
40	7	0.9	194	4	US-09-489-352-17	Sequence 17, Appl
41	7	0.9	195	4	US-09-117-257-34	Sequence 34, Appl
42	7	0.9	195	4	US-09-117-257-46	Sequence 46, Appl
43	7	0.9	195	4	US-09-489-352-34	Sequence 34, Appl
44	7	0.9	195	4	US-09-489-352-46	Sequence 46, Appl
45	7	0.9	198	4	US-09-228-986-93	Sequence 93, Appl
46	7	0.9	327	1	US-08-211-682-25	Sequence 25, Appl
47	7	0.9	374	4	US-09-091-405-2	Sequence 2, Appl
48	7	0.9	377	1	US-08-525-697-2	Sequence 2, Appl
49	7	0.9	396	2	US-08-775-428-2	Sequence 2, Appl
50	7	0.9	474	4	US-09-461-474-10	Sequence 10, Appl
51	7	0.9	512	2	US-08-194-981E-5	Sequence 5, Appl
52	7	0.9	627	1	US-08-703-947-2	Sequence 2, Appl
53	7	0.9	631	4	US-09-147-119-7	Sequence 7, Appl
54	7	0.9	659	1	US-08-258-639A-4	Sequence 4, Appl
55	7	0.9	659	1	US-08-900-951-4	Sequence 4, Appl
56	7	0.9	659	5	PCT-US95-07391A-4	Sequence 4, Appl
57	7	0.9	661	1	US-08-514-014-4	Sequence 4, Appl
58	7	0.9	661	2	US-08-833-823-4	Sequence 4, Appl
59	7	0.9	677	3	US-08-480-640A-115	Sequence 115, App
60	7	0.9	677	3	US-08-480-640A-193	Sequence 193, App
61	7	0.9	677	3	US-08-295-802-115	Sequence 115, App
62	7	0.9	677	4	US-08-686-968C-58	Sequence 58, Appl
63	7	0.9	677	4	US-08-686-968C-193	Sequence 193, App
64	7	0.9	677	4	US-08-488-237A-115	Sequence 115, App
65	7	0.9	677	4	US-08-488-237A-193	Sequence 193, App
66	7	0.9	677	4	US-08-375-992A-115	Sequence 115, App
67	7	0.9	677	4	US-08-375-992A-193	Sequence 193, App
68	7	0.9	711	4	US-09-228-986-79	Sequence 79, Appl
69	7	0.9	726	4	US-09-413-814-30	Sequence 30, Appl
70	7	0.9	771	4	US-09-188-930-183	Sequence 183, App
71	7	0.9	807	4	US-09-081-345-2	Sequence 2, Appl
72	7	0.9	915	2	US-08-480-917-2	Sequence 2, Appl
73	7	0.9	915	4	US-09-138-736-2	Sequence 2, Appl
74	7	0.9	999	2	US-08-473-553A-5	Sequence 5, Appl
75	7	0.9	1142	2	US-08-993-118-7	Sequence 7, Appl
76	7	0.9	1142	3	US-08-845-528C-7	Sequence 7, Appl
77	7	0.9	1142	4	US-09-061-709-2	Sequence 2, Appl
78	7	0.9	1196	4	US-08-881-706-2	Sequence 2, Appl
79	7	0.9	1525	3	US-09-191-647-2	Sequence 2, Appl
80	7	0.9	1525	4	US-09-540-245A-2	Sequence 2, Appl
81	7	0.9	1525	4	US-09-540-153-2	Sequence 2, Appl
82	7	0.9	1817	4	US-09-004-838-125	Sequence 125, App
83	7	0.9	2343	4	US-09-324-867-2	Sequence 2, Appl
84	6	0.7	10	3	US-08-898-351-3	Sequence 3, Appl
85	6	0.7	11	4	US-08-652-877-72	Sequence 72, Appl
86	6	0.7	11	4	US-08-476-515A-72	Sequence 72, Appl
87	6	0.7	11	4	US-09-025-596-2	Sequence 2, Appl
88	6	0.7	16	2	US-08-480-190-41	Sequence 41, Appl
89	6	0.7	16	2	US-08-488-379-41	Sequence 41, Appl
90	6	0.7	16	5	PCT-US93-07545-41	Sequence 41, Appl
91	6	0.7	17	2	US-08-468-279-2	Sequence 2, Appl
92	6	0.7	20	1	US-08-080-809-3	Sequence 3, Appl
93	6	0.7	20	2	US-08-484-530-47	Sequence 47, Appl
94	6	0.7	20	2	US-08-827-618A-47	Sequence 47, Appl
95	6	0.7	20	3	US-08-483-952A-47	Sequence 47, Appl
96	6	0.7	20	3	US-08-855-925A-2	Sequence 2, Appl
97	6	0.7	20	4	US-08-973-782-1	Sequence 1, Appl
98	6	0.7	20	5	PCT-US96-11374-1	Sequence 1, Appl
99	6	0.7	21	4	US-08-905-223-20	Sequence 20, Appl
100	6	0.7	21	4	US-09-247-155-20	Sequence 20, Appl

101	6	0.7	24	3	US-08-592-500-26	Sequence 26, Appl	174	6	0.7	111	1	US-08-411-796-218	Sequence 218, Appl
102	6	0.7	24	3	US-08-195-006-26	Sequence 26, Appl	175	6	0.7	111	2	US-08-469-319A-65	Sequence 65, Appl
103	6	0.7	24	5	PCT-US94-07644A-26	Sequence 26, Appl	176	6	0.7	111	2	US-08-469-319A-69	Sequence 69, Appl
104	6	0.7	30	3	US-08-248-839C-80	Sequence 80, Appl	177	6	0.7	111	2	US-08-469-319A-70	Sequence 70, Appl
105	6	0.7	34	3	US-08-658-136-33	Sequence 33, Appl	178	6	0.7	111	2	US-08-469-319A-71	Sequence 71, Appl
106	6	0.7	40	4	US-08-905-223-387	Sequence 387, App	179	6	0.7	111	2	US-08-469-319A-72	Sequence 72, Appl
107	6	0.7	42	1	US-08-340-812-7	Sequence 7, Appl	180	6	0.7	111	2	US-08-469-319A-73	Sequence 73, Appl
108	6	0.7	42	1	US-08-459-064B-7	Sequence 7, Appl	181	6	0.7	111	2	US-08-469-319A-74	Sequence 74, Appl
109	6	0.7	42	5	PCT-US93-00909-7	Sequence 7, Appl	182	6	0.7	111	2	US-08-469-319A-75	Sequence 75, Appl
110	6	0.7	42	5	PCT-US93-00909-7	Sequence 7, Appl	183	6	0.7	111	2	US-08-469-319A-76	Sequence 76, Appl
111	6	0.7	47	4	US-09-227-357-601	Sequence 601, App	184	6	0.7	111	2	US-08-469-319A-77	Sequence 77, Appl
112	6	0.7	52	1	US-08-081-539-27	Sequence 27, Appl	185	6	0.7	111	2	US-08-469-319A-78	Sequence 78, Appl
113	6	0.7	52	1	US-08-466-647-27	Sequence 27, Appl	186	6	0.7	111	2	US-08-469-319A-407	Sequence 407, App
114	6	0.7	52	1	US-08-411-795B-146	Sequence 146, App	187	6	0.7	111	3	US-08-471-039-23	Sequence 23, Appl
115	6	0.7	52	1	US-08-411-796-146	Sequence 146, App	188	6	0.7	111	3	US-08-471-039-65	Sequence 65, Appl
116	6	0.7	52	2	US-08-469-319A-146	Sequence 146, App	189	6	0.7	111	3	US-08-471-039-217	Sequence 217, App
117	6	0.7	52	3	US-08-471-039-146	Sequence 146, App	190	6	0.7	111	3	US-08-471-039-218	Sequence 218, App
118	6	0.7	52	5	PCT-US93-11198-146	Sequence 146, App	191	6	0.7	111	3	US-08-469-318-12	Sequence 12, Appl
119	6	0.7	55	4	US-09-227-357-553	Sequence 553, App	192	6	0.7	111	3	US-08-469-318-13	Sequence 13, Appl
120	6	0.7	66	4	US-09-353-585-7	Sequence 7, Appl	193	6	0.7	111	3	US-08-469-318-14	Sequence 14, Appl
121	6	0.7	70	3	US-08-989-510A-5	Sequence 5, Appl	194	6	0.7	111	3	US-08-469-318-15	Sequence 15, Appl
122	6	0.7	70	3	US-08-989-510A-19	Sequence 19, Appl	195	6	0.7	111	3	US-08-469-318-16	Sequence 16, Appl
123	6	0.7	70	4	US-09-182-816-5	Sequence 5, Appl	196	6	0.7	111	3	US-08-469-318-17	Sequence 17, Appl
124	6	0.7	70	4	US-09-182-816-20	Sequence 20, Appl	197	6	0.7	111	3	US-08-469-318-18	Sequence 18, Appl
125	6	0.7	70	4	US-09-471-528-5	Sequence 5, Appl	198	6	0.7	111	3	US-08-469-318-19	Sequence 19, Appl
126	6	0.7	70	4	US-09-471-528-20	Sequence 20, Appl	199	6	0.7	111	3	US-08-469-318-20	Sequence 20, Appl
127	6	0.7	70	4	US-09-634-530-5	Sequence 5, Appl	200	6	0.7	111	3	US-08-469-318-21	Sequence 21, Appl
128	6	0.7	70	4	US-09-634-530-20	Sequence 20, Appl	201	6	0.7	111	3	US-08-468-609A-12	Sequence 12, Appl
129	6	0.7	75	1	US-08-081-539-23	Sequence 23, Appl	202	6	0.7	111	3	US-08-468-609A-13	Sequence 13, Appl
130	6	0.7	75	1	US-08-466-647-23	Sequence 23, Appl	203	6	0.7	111	3	US-08-468-609A-14	Sequence 14, Appl
131	6	0.7	75	1	US-08-411-795B-139	Sequence 139, App	204	6	0.7	111	3	US-08-468-609A-15	Sequence 15, Appl
132	6	0.7	75	1	US-08-665-220-66	Sequence 66, Appl	205	6	0.7	111	3	US-08-468-609A-16	Sequence 16, Appl
133	6	0.7	75	2	US-08-469-319A-139	Sequence 139, App	206	6	0.7	111	3	US-08-468-609A-17	Sequence 17, Appl
134	6	0.7	75	4	US-09-291-692-66	Sequence 66, Appl	207	6	0.7	111	3	US-08-468-609A-18	Sequence 18, Appl
135	6	0.7	76	4	US-09-238-373-4	Sequence 4, Appl	208	6	0.7	111	3	US-08-468-609A-19	Sequence 19, Appl
136	6	0.7	77	4	US-09-227-357-552	Sequence 552, App	209	6	0.7	111	3	US-08-468-609A-20	Sequence 20, Appl
137	6	0.7	81	2	US-08-807-200-4	Sequence 4, Appl	210	6	0.7	111	3	US-08-468-609A-21	Sequence 21, Appl
138	6	0.7	81	4	US-09-001-777-4	Sequence 4, Appl	211	6	0.7	111	3	US-08-192-325B-12	Sequence 12, Appl
139	6	0.7	83	4	US-09-382-155-11	Sequence 11, Appl	212	6	0.7	111	3	US-08-192-325B-13	Sequence 13, Appl
140	6	0.7	83	4	US-09-074-044A-11	Sequence 11, Appl	213	6	0.7	111	3	US-08-192-325B-14	Sequence 14, Appl
141	6	0.7	86	2	US-08-245-511-38	Sequence 38, Appl	214	6	0.7	111	3	US-08-192-325B-15	Sequence 15, Appl
142	6	0.7	86	2	US-08-600-993A-38	Sequence 38, Appl	215	6	0.7	111	3	US-08-192-325B-16	Sequence 16, Appl
143	6	0.7	96	1	US-08-486-013-21	Sequence 21, Appl	216	6	0.7	111	3	US-08-192-325B-17	Sequence 17, Appl
144	6	0.7	96	2	US-08-482-279-21	Sequence 21, Appl	217	6	0.7	111	3	US-08-192-325B-18	Sequence 18, Appl
145	6	0.7	96	2	US-08-342-268-21	Sequence 21, Appl	218	6	0.7	111	3	US-08-192-325B-19	Sequence 19, Appl
146	6	0.7	96	3	US-09-015-968-21	Sequence 21, Appl	219	6	0.7	111	3	US-08-192-325B-20	Sequence 20, Appl
147	6	0.7	96	4	US-09-397-386-21	Sequence 21, Appl	220	6	0.7	111	3	US-08-192-325B-21	Sequence 21, Appl
148	6	0.7	104	2	US-08-853-97A-3	Sequence 3, Appl	221	6	0.7	111	3	US-08-471-045-12	Sequence 12, Appl
149	6	0.7	104	4	US-09-172-988-3	Sequence 3, Appl	222	6	0.7	111	3	US-08-471-045-13	Sequence 13, Appl
150	6	0.7	105	1	US-08-081-539-73	Sequence 73, Appl	223	6	0.7	111	3	US-08-471-045-14	Sequence 14, Appl
151	6	0.7	105	1	US-08-466-647-73	Sequence 73, Appl	224	6	0.7	111	3	US-08-471-045-15	Sequence 15, Appl
152	6	0.7	106	4	US-09-188-930-136	Sequence 136, App	225	6	0.7	111	3	US-08-471-045-16	Sequence 16, Appl
153	6	0.7	107	1	US-08-369-796-17	Sequence 17, Appl	226	6	0.7	111	3	US-08-471-045-17	Sequence 17, Appl
154	6	0.7	107	2	US-08-852-091-17	Sequence 17, Appl	227	6	0.7	111	3	US-08-471-045-18	Sequence 18, Appl
155	6	0.7	107	3	US-08-478-097A-19	Sequence 19, Appl	228	6	0.7	111	3	US-08-471-045-19	Sequence 19, Appl
156	6	0.7	107	5	PCT-US95-17025-17	Sequence 17, Appl	229	6	0.7	111	3	US-08-471-045-20	Sequence 20, Appl
157	6	0.7	110	1	US-08-081-539-69	Sequence 69, Appl	230	6	0.7	111	3	US-08-471-045-21	Sequence 21, Appl
158	6	0.7	110	1	US-08-466-647-69	Sequence 69, Appl	231	6	0.7	111	3	US-08-469-712A-12	Sequence 12, Appl
159	6	0.7	111	1	US-08-411-795B-65	Sequence 65, Appl	232	6	0.7	111	3	US-08-469-712A-13	Sequence 13, Appl
160	6	0.7	111	1	US-08-411-795B-69	Sequence 69, Appl	233	6	0.7	111	3	US-08-469-712A-14	Sequence 14, Appl
161	6	0.7	111	1	US-08-411-795B-70	Sequence 70, Appl	234	6	0.7	111	3	US-08-469-712A-15	Sequence 15, Appl
162	6	0.7	111	1	US-08-411-795B-71	Sequence 71, Appl	235	6	0.7	111	3	US-08-469-712A-16	Sequence 16, Appl
163	6	0.7	111	1	US-08-411-795B-72	Sequence 72, Appl	236	6	0.7	111	3	US-08-469-712A-17	Sequence 17, Appl
164	6	0.7	111	1	US-08-411-795B-73	Sequence 73, Appl	237	6	0.7	111	3	US-08-469-712A-18	Sequence 18, Appl
165	6	0.7	111	1	US-08-411-795B-74	Sequence 74, Appl	238	6	0.7	111	3	US-08-469-712A-19	Sequence 19, Appl
166	6	0.7	111	1	US-08-411-795B-75	Sequence 75, Appl	239	6	0.7	111	3	US-08-469-712A-20	Sequence 20, Appl
167	6	0.7	111	1	US-08-411-795B-76	Sequence 76, Appl	240	6	0.7	111	3	US-08-469-712A-21	Sequence 21, Appl
168	6	0.7	111	1	US-08-411-795B-77	Sequence 77, Appl	241	6	0.7	111	4	US-08-193-373A-12	Sequence 12, Appl
169	6	0.7	111	1	US-08-411-795B-78	Sequence 78, Appl	242	6	0.7	111	4	US-08-193-373A-13	Sequence 13, Appl
170	6	0.7	111	1	US-08-411-795B-407	Sequence 407, App	243	6	0.7	111	4	US-08-193-373A-14	Sequence 14, Appl
171	6	0.7	111	1	US-08-411-796-23	Sequence 23, Appl	244	6	0.7	111	4	US-08-193-373A-15	Sequence 15, Appl
172	6	0.7	111	1	US-08-411-796-65	Sequence 65, Appl	245	6	0.7	111	4	US-08-193-373A-16	Sequence 16, Appl
173	6	0.7	111	1	US-08-411-796-217	Sequence 217, App	246	6	0.7	111	4	US-08-193-373A-17	Sequence 17, Appl



247	6	0.7	111	4	US-08-193-373A-18	Sequence 18, Appl	320	6	0.7	113	3	US-08-468-609A-25	Sequence 25, Appl
248	6	0.7	111	4	US-08-193-373A-19	Sequence 19, Appl	321	6	0.7	113	3	US-08-468-609A-26	Sequence 26, Appl
249	6	0.7	111	4	US-08-193-373A-20	Sequence 20, Appl	322	6	0.7	113	3	US-08-468-609A-27	Sequence 27, Appl
250	6	0.7	111	4	US-08-193-373A-21	Sequence 21, Appl	323	6	0.7	113	3	US-08-192-325B-25	Sequence 25, Appl
251	6	0.7	111	5	PCT-US93-11198-21	Sequence 23, Appl	324	6	0.7	113	3	US-08-192-325B-26	Sequence 26, Appl
252	6	0.7	111	5	PCT-US93-11198-22	Sequence 65, Appl	325	6	0.7	113	3	US-08-192-325B-27	Sequence 27, Appl
253	6	0.7	111	5	PCT-US93-11198-217	Sequence 218, App	326	6	0.7	113	3	US-08-471-045-25	Sequence 25, Appl
254	6	0.7	111	5	PCT-US93-11198-218	Sequence 217, App	327	6	0.7	113	3	US-08-471-045-26	Sequence 26, Appl
255	6	0.7	111	5	PCT-US95-01185-12	Sequence 12, Appl	328	6	0.7	113	3	US-08-471-045-27	Sequence 27, Appl
256	6	0.7	111	5	PCT-US95-01185-13	Sequence 13, Appl	329	6	0.7	113	3	US-08-469-712A-25	Sequence 25, Appl
257	6	0.7	111	5	PCT-US95-01185-14	Sequence 14, Appl	330	6	0.7	113	3	US-08-469-712A-26	Sequence 26, Appl
258	6	0.7	111	5	PCT-US95-01185-15	Sequence 15, Appl	331	6	0.7	113	3	US-08-469-712A-27	Sequence 27, Appl
259	6	0.7	111	5	PCT-US95-01185-16	Sequence 16, Appl	332	6	0.7	113	4	US-08-193-373A-25	Sequence 25, Appl
260	6	0.7	111	5	PCT-US95-01185-17	Sequence 17, Appl	333	6	0.7	113	4	US-08-193-373A-26	Sequence 26, Appl
261	6	0.7	111	5	PCT-US95-01185-18	Sequence 18, Appl	334	6	0.7	113	4	US-08-193-373A-27	Sequence 27, Appl
262	6	0.7	111	5	PCT-US95-01185-19	Sequence 19, Appl	335	6	0.7	113	5	PCT-US93-11198-13	Sequence 13, Appl
263	6	0.7	111	5	PCT-US95-01185-20	Sequence 20, Appl	336	6	0.7	113	5	PCT-US93-11198-66	Sequence 66, Appl
264	6	0.7	111	5	PCT-US95-01185-21	Sequence 21, Appl	337	6	0.7	113	5	PCT-US93-11198-67	Sequence 67, Appl
265	6	0.7	112	1	US-08-081-539-18	Sequence 18, Appl	338	6	0.7	113	5	PCT-US93-11198-69	Sequence 69, Appl
266	6	0.7	112	1	US-08-081-539-19	Sequence 19, Appl	339	6	0.7	113	5	PCT-US95-01185-25	Sequence 25, Appl
267	6	0.7	112	1	US-08-081-539-65	Sequence 65, Appl	340	6	0.7	113	5	PCT-US95-01185-26	Sequence 26, Appl
268	6	0.7	112	1	US-08-081-539-66	Sequence 66, Appl	341	6	0.7	113	5	PCT-US95-01185-27	Sequence 27, Appl
269	6	0.7	112	1	US-08-081-539-67	Sequence 67, Appl	342	6	0.7	114	1	US-08-081-539-64	Sequence 64, Appl
270	6	0.7	112	1	US-08-081-539-68	Sequence 68, Appl	343	6	0.7	114	1	US-08-466-647-64	Sequence 64, Appl
271	6	0.7	112	1	US-08-081-539-70	Sequence 70, Appl	344	6	0.7	114	4	US-09-627-376-17	Sequence 17, Appl
272	6	0.7	112	1	US-08-081-539-71	Sequence 71, Appl	345	6	0.7	120	1	US-08-081-539-10	Sequence 10, Appl
273	6	0.7	112	1	US-08-081-539-72	Sequence 72, Appl	346	6	0.7	120	1	US-08-081-539-14	Sequence 14, Appl
274	6	0.7	112	1	US-08-466-647-18	Sequence 18, Appl	347	6	0.7	120	1	US-08-466-647-10	Sequence 10, Appl
275	6	0.7	112	1	US-08-466-647-19	Sequence 19, Appl	348	6	0.7	120	1	US-08-466-647-14	Sequence 14, Appl
276	6	0.7	112	1	US-08-466-647-65	Sequence 65, Appl	349	6	0.7	120	1	US-08-466-647-14	Sequence 14, Appl
277	6	0.7	112	1	US-08-466-647-66	Sequence 66, Appl	350	6	0.7	120	1	US-08-411-795B-135	Sequence 135, App
278	6	0.7	112	1	US-08-466-647-67	Sequence 67, Appl	351	6	0.7	120	1	US-08-411-795B-402	Sequence 402, App
279	6	0.7	112	1	US-08-466-647-68	Sequence 68, Appl	352	6	0.7	120	2	US-08-469-319A-135	Sequence 135, App
280	6	0.7	112	1	US-08-466-647-70	Sequence 70, Appl	353	6	0.7	120	2	US-08-469-319A-402	Sequence 402, App
281	6	0.7	112	1	US-08-466-647-71	Sequence 71, Appl	354	6	0.7	120	3	US-08-471-039-135	Sequence 135, App
282	6	0.7	112	1	US-08-466-647-72	Sequence 72, Appl	355	6	0.7	120	5	PCT-US93-11198-135	Sequence 135, App
283	6	0.7	112	1	US-08-411-795B-138	Sequence 138, App	356	6	0.7	124	1	US-08-455-559-11	Sequence 11, Appl
284	6	0.7	112	2	US-08-469-319A-138	Sequence 138, App	357	6	0.7	124	4	US-09-145-060-11	Sequence 11, Appl
285	6	0.7	113	1	US-08-081-539-20	Sequence 20, Appl	358	6	0.7	124	5	PCT-US94-00657-11	Sequence 11, Appl
286	6	0.7	113	1	US-08-081-539-76	Sequence 76, Appl	359	6	0.7	126	1	US-08-081-539-17	Sequence 17, Appl
287	6	0.7	113	1	US-08-466-647-20	Sequence 20, Appl	360	6	0.7	126	1	US-08-466-647-17	Sequence 17, Appl
288	6	0.7	113	1	US-08-466-647-76	Sequence 76, Appl	361	6	0.7	126	1	US-08-411-795B-10	Sequence 10, Appl
289	6	0.7	113	1	US-08-411-795B-13	Sequence 13, Appl	362	6	0.7	126	1	US-08-411-795B-10	Sequence 10, Appl
290	6	0.7	113	1	US-08-411-795B-82	Sequence 82, Appl	363	6	0.7	126	2	US-08-469-319A-10	Sequence 10, Appl
291	6	0.7	113	1	US-08-411-795B-83	Sequence 83, Appl	364	6	0.7	126	3	US-08-471-039-10	Sequence 10, Appl
292	6	0.7	113	1	US-08-411-795B-84	Sequence 84, Appl	365	6	0.7	126	5	PCT-US93-11198-10	Sequence 10, Appl
293	6	0.7	113	1	US-08-411-795B-259	Sequence 259, App	366	6	0.7	131	6	5166322-6	Patent No. 5166322
294	6	0.7	113	1	US-08-411-795B-260	Sequence 260, App	367	6	0.7	133	1	US-08-318-193-53	Sequence 53, Appl
295	6	0.7	113	1	US-08-411-795B-261	Sequence 261, App	368	6	0.7	133	6	5166322-4	Patent No. 5166322
296	6	0.7	113	1	US-08-411-795B-262	Sequence 262, App	369	6	0.7	134	1	US-08-081-539-13	Sequence 13, Appl
297	6	0.7	113	1	US-08-411-795B-263	Sequence 263, App	370	6	0.7	134	1	US-08-466-647-13	Sequence 13, Appl
298	6	0.7	113	1	US-08-411-795B-406	Sequence 406, App	371	6	0.7	134	1	US-08-411-795B-128	Sequence 128, App
299	6	0.7	113	1	US-08-411-796-13	Sequence 13, Appl	372	6	0.7	134	1	US-08-411-796-128	Sequence 128, App
300	6	0.7	113	1	US-08-411-796-66	Sequence 66, Appl	373	6	0.7	134	1	US-08-192-299B-9	Sequence 9, Appl
301	6	0.7	113	1	US-08-411-796-67	Sequence 67, Appl	374	6	0.7	134	1	US-08-191-973B-9	Sequence 9, Appl
302	6	0.7	113	1	US-08-411-796-69	Sequence 69, Appl	375	6	0.7	134	2	US-08-469-319A-128	Sequence 128, App
303	6	0.7	113	2	US-08-469-319A-13	Sequence 13, Appl	376	6	0.7	134	2	US-08-470-775-9	Sequence 9, Appl
304	6	0.7	113	2	US-08-469-319A-82	Sequence 82, Appl	377	6	0.7	134	2	US-08-470-509-9	Sequence 9, Appl
305	6	0.7	113	2	US-08-469-319A-83	Sequence 83, Appl	378	6	0.7	134	2	US-08-559-009-9	Sequence 9, Appl
306	6	0.7	113	2	US-08-469-319A-84	Sequence 84, Appl	379	6	0.7	134	3	US-08-471-039-128	Sequence 128, App
307	6	0.7	113	2	US-08-469-319A-259	Sequence 259, App	380	6	0.7	134	3	US-08-469-318-49	Sequence 49, Appl
308	6	0.7	113	2	US-08-469-319A-260	Sequence 260, App	381	6	0.7	134	3	US-08-468-609A-49	Sequence 49, Appl
309	6	0.7	113	2	US-08-469-319A-261	Sequence 261, App	382	6	0.7	134	3	US-08-192-325B-49	Sequence 49, Appl
310	6	0.7	113	2	US-08-469-319A-262	Sequence 262, App	383	6	0.7	134	3	US-08-471-045-49	Sequence 49, Appl
311	6	0.7	113	2	US-08-469-319A-263	Sequence 263, App	384	6	0.7	134	3	US-08-559-267A-9	Sequence 9, Appl
312	6	0.7	113	2	US-08-469-319A-406	Sequence 406, App	385	6	0.7	134	3	US-08-469-712A-49	Sequence 49, Appl
313	6	0.7	113	3	US-08-471-039-13	Sequence 13, Appl	386	6	0.7	134	4	US-08-469-124-9	Sequence 9, Appl
314	6	0.7	113	3	US-08-471-039-66	Sequence 66, Appl	387	6	0.7	134	4	US-09-068-655-12	Sequence 12, Appl
315	6	0.7	113	3	US-08-471-039-67	Sequence 67, Appl	388	6	0.7	134	4	US-08-193-373A-49	Sequence 49, Appl
316	6	0.7	113	3	US-08-471-039-69	Sequence 69, Appl	389	6	0.7	134	5	PCT-US93-11198-128	Sequence 128, App
317	6	0.7	113	3	US-08-469-318-25	Sequence 25, Appl	390	6	0.7	134	5	PCT-US94-04208-4	Sequence 4, Appl
318	6	0.7	113	3	US-08-469-318-26	Sequence 26, Appl	391	6	0.7	134	5	PCT-US95-01185-49	Sequence 49, Appl
319	6	0.7	113	3	US-08-469-318-27	Sequence 27, Appl	392	6	0.7	134	6	5166322-2	Patent No. 5166322

393	6	0.7	142	4	US-08-945-983-7	Sequence 7, Appl1	466	6	0.7	213	3	US-08-935-333-3	Sequence 3, Appl1
394	6	0.7	145	1	US-08-150-331-46	Sequence 46, Appl	467	6	0.7	214	1	US-07-739-642-6	Sequence 6, Appl1
395	6	0.7	149	1	US-08-530-010-15	Sequence 15, Appl	468	6	0.7	214	1	US-07-739-643-12	Sequence 12, Appl
396	6	0.7	149	2	US-08-484-101B-15	Sequence 15, Appl	469	6	0.7	214	1	US-07-739-643-6	Sequence 6, Appl1
397	6	0.7	149	4	US-08-714-524D-15	Sequence 15, Appl	470	6	0.7	214	1	US-07-739-643-12	Sequence 12, Appl
398	6	0.7	150	2	US-08-387-942C-52	Sequence 52, Appl	471	6	0.7	214	1	US-07-739-142-6	Sequence 6, Appl1
399	6	0.7	150	4	US-09-188-930-306	Sequence 306, App	472	6	0.7	214	1	US-07-739-142-12	Sequence 12, Appl
400	6	0.7	150	6	5304637-1	Patent No. 5304637	473	6	0.7	220	2	US-08-807-200-2	Sequence 2, Appl1
401	6	0.7	151	2	US-08-387-942C-46	Sequence 46, Appl	474	6	0.7	220	4	US-09-001-777-2	Sequence 2, Appl1
402	6	0.7	152	1	US-08-284-393B-12	Sequence 12, Appl	475	6	0.7	224	5	PCT-US91-09055-4	Sequence 4, Appl1
403	6	0.7	152	4	US-08-470-369-2	Sequence 2, Appl1	476	6	0.7	226	1	US-07-929-198-4	Sequence 4, Appl1
404	6	0.7	152	5	PCT-US95-08950-12	Sequence 12, Appl	477	6	0.7	226	3	US-09-176-657-3	Sequence 3, Appl1
405	6	0.7	154	3	US-08-968-747-1	Sequence 1, Appl1	478	6	0.7	231	4	US-08-974-380-2	Sequence 2, Appl1
406	6	0.7	154	4	US-09-228-986-84	Sequence 84, Appl	479	6	0.7	234	3	US-08-836-236-7	Sequence 7, Appl1
407	6	0.7	155	3	US-08-968-747-17	Sequence 17, Appl	480	6	0.7	235	2	US-08-924-759-10	Sequence 10, Appl
408	6	0.7	155	4	US-08-432-994A-8	Sequence 8, Appl1	481	6	0.7	235	3	US-09-248-335-10	Sequence 5, Appl1
409	6	0.7	159	4	US-08-445-585-3	Sequence 3, Appl1	482	6	0.7	235	5	PCT-US96-10521-5	Sequence 22, Appl
410	6	0.7	161	3	US-08-968-747-19	Sequence 19, Appl	483	6	0.7	236	2	US-08-464-517-22	Sequence 22, Appl
411	6	0.7	166	2	US-08-775-978-1	Sequence 1, Appl1	484	6	0.7	236	3	US-08-246-361A-22	Sequence 22, Appl
412	6	0.7	166	2	US-08-775-978-3	Sequence 3, Appl1	485	6	0.7	236	3	US-08-463-772-22	Sequence 22, Appl
413	6	0.7	168	4	US-09-087-465-32	Sequence 32, Appl	486	6	0.7	236	5	PCT-US93-05000-22	Sequence 22, Appl
414	6	0.7	180	4	US-09-382-155-18	Sequence 18, Appl	487	6	0.7	238	1	US-08-928-443-5	Sequence 5, Appl1
415	6	0.7	180	4	US-09-074-044A-18	Sequence 18, Appl	488	6	0.7	238	3	US-09-129-053-5	Sequence 5, Appl1
416	6	0.7	181	2	US-08-932-142-2	Sequence 2, Appl1	489	6	0.7	240	1	US-08-261-822A-80	Sequence 80, Appl
417	6	0.7	181	4	US-09-342-458-2	Sequence 2, Appl1	490	6	0.7	240	5	PCT-US95-07744A-80	Sequence 80, Appl
418	6	0.7	182	1	US-08-466-603-5	Sequence 5, Appl1	491	6	0.7	241	1	US-08-188-228-56	Sequence 56, Appl
419	6	0.7	182	1	US-08-314-503A-5	Sequence 5, Appl1	492	6	0.7	241	1	US-08-332-643-50	Sequence 50, Appl
420	6	0.7	182	1	US-08-468-065-5	Sequence 5, Appl1	493	6	0.7	241	1	US-08-332-638-56	Sequence 56, Appl
421	6	0.7	182	2	US-08-466-717-5	Sequence 5, Appl1	494	6	0.7	247	4	US-09-228-986-105	Sequence 105, App
422	6	0.7	182	3	US-08-466-743-5	Sequence 5, Appl1	495	6	0.7	249	1	US-08-466-603-2	Sequence 2, Appl1
423	6	0.7	182	5	PCT-US95-12414-5	Sequence 5, Appl1	496	6	0.7	249	1	US-08-314-503A-2	Sequence 2, Appl1
424	6	0.7	183	3	US-08-968-747-20	Sequence 20, Appl	497	6	0.7	249	1	US-08-468-066-2	Sequence 2, Appl1
425	6	0.7	183	4	US-09-248-588-2	Sequence 2, Appl1	498	6	0.7	249	2	US-08-466-717-2	Sequence 2, Appl1
426	6	0.7	183	4	US-09-248-588-4	Sequence 4, Appl1	499	6	0.7	249	2	US-08-766-738-4	Sequence 4, Appl1
427	6	0.7	183	5	PCT-US96-10602-12	Sequence 12, Appl	500	6	0.7	249	3	US-08-466-743-2	Sequence 2, Appl1
428	6	0.7	184	4	US-09-117-257-21	Sequence 21, Appl	501	6	0.7	249	5	PCT-US95-12414-2	Sequence 2, Appl1
429	6	0.7	184	4	US-08-945-476-21	Sequence 21, Appl	502	6	0.7	251	2	US-08-766-738-1	Sequence 1, Appl1
430	6	0.7	184	4	US-09-489-352-21	Sequence 21, Appl	503	6	0.7	251	2	US-08-766-738-3	Sequence 3, Appl1
431	6	0.7	185	1	US-07-739-642-2	Sequence 2, Appl1	504	6	0.7	252	3	US-08-906-769-103	Sequence 103, App
432	6	0.7	185	1	US-07-739-643-4	Sequence 4, Appl1	505	6	0.7	252	3	US-08-906-616-103	Sequence 103, App
433	6	0.7	185	1	US-07-739-642-8	Sequence 8, Appl1	506	6	0.7	252	4	US-08-817-795-103	Sequence 103, App
434	6	0.7	185	1	US-07-739-642-10	Sequence 10, Appl	507	6	0.7	252	4	US-08-639-075A-103	Sequence 103, App
435	6	0.7	185	1	US-07-739-643-2	Sequence 2, Appl1	508	6	0.7	252	4	US-09-012-431-103	Sequence 103, App
436	6	0.7	185	1	US-07-739-643-4	Sequence 4, Appl1	509	6	0.7	252	4	US-09-012-692-103	Sequence 103, App
437	6	0.7	185	1	US-07-739-643-8	Sequence 8, Appl1	510	6	0.7	252	4	US-08-906-613-103	Sequence 103, App
438	6	0.7	185	1	US-07-739-643-10	Sequence 10, Appl	511	6	0.7	252	5	PCT-US95-14442A-103	Sequence 103, App
439	6	0.7	185	1	US-07-739-142-2	Sequence 2, Appl1	512	6	0.7	257	1	US-08-618-164-2	Sequence 2, Appl1
440	6	0.7	185	1	US-07-739-142-4	Sequence 4, Appl1	513	6	0.7	257	3	US-08-728-603-19	Sequence 19, Appl
441	6	0.7	185	1	US-07-739-142-8	Sequence 8, Appl1	514	6	0.7	258	1	US-07-990-301A-2	Sequence 2, Appl1
442	6	0.7	185	1	US-07-739-142-10	Sequence 10, Appl	515	6	0.7	261	5	PCT-US96-10521-25	Sequence 25, Appl
443	6	0.7	185	4	US-09-117-257-52	Sequence 52, Appl	516	6	0.7	271	4	US-09-323-427-9	Sequence 9, Appl1
444	6	0.7	185	4	US-09-248-588-6	Sequence 6, Appl1	517	6	0.7	274	2	US-08-867-030B-15	Sequence 15, Appl
445	6	0.7	185	4	US-09-489-352-52	Sequence 52, Appl	518	6	0.7	274	5	PCT-US95-06119-15	Sequence 15, Appl
446	6	0.7	188	4	US-09-068-655-5	Sequence 5, Appl1	519	6	0.7	275	4	US-09-147-915-4	Sequence 4, Appl1
447	6	0.7	189	2	US-08-464-517-21	Sequence 21, Appl	520	6	0.7	277	1	US-08-118-270-62	Sequence 62, Appl
448	6	0.7	189	2	US-08-246-361A-21	Sequence 21, Appl	521	6	0.7	277	2	US-08-567-375-16	Sequence 16, Appl
449	6	0.7	189	3	US-08-463-772-21	Sequence 21, Appl	522	6	0.7	277	5	PCT-US93-08528-62	Sequence 62, Appl
450	6	0.7	189	5	PCT-US93-05000-21	Sequence 21, Appl	523	6	0.7	277	5	PCT-US93-08528-62	Sequence 62, Appl
451	6	0.7	191	2	US-08-286-819A-21	Sequence 21, Appl	524	6	0.7	283	4	US-09-434-774-6	Sequence 6, Appl1
452	6	0.7	191	3	US-08-980-357-21	Sequence 21, Appl	525	6	0.7	284	2	US-08-786-606-9	Sequence 9, Appl1
453	6	0.7	191	4	US-08-842-306B-6	Sequence 6, Appl1	526	6	0.7	284	4	US-09-069-023-5	Sequence 5, Appl1
454	6	0.7	191	4	US-08-838-973B-6	Sequence 6, Appl1	527	6	0.7	289	2	US-08-246-361A-4	Sequence 4, Appl1
455	6	0.7	193	3	US-08-968-747-2	Sequence 2, Appl1	528	6	0.7	289	5	PCT-US93-05000-4	Sequence 4, Appl1
456	6	0.7	194	3	US-08-968-747-18	Sequence 18, Appl	529	6	0.7	289	5	PCT-US96-10602-8	Sequence 8, Appl1
457	6	0.7	195	1	US-08-236-427-4	Sequence 4, Appl1	530	6	0.7	293	3	US-08-812-586-3	Sequence 3, Appl1
458	6	0.7	196	4	US-09-383-586-11	Sequence 11, Appl	531	6	0.7	294	4	US-09-242-948-2	Sequence 2, Appl1
459	6	0.7	199	3	US-08-968-747-21	Sequence 21, Appl	532	6	0.7	300	4	US-09-099-041A-4	Sequence 4, Appl1
460	6	0.7	206	4	US-09-228-986-120	Sequence 120, App	533	6	0.7	301	1	US-07-797-553-2	Sequence 2, Appl1
461	6	0.7	210	2	US-08-416-603-8	Sequence 8, Appl1	534	6	0.7	301	4	US-09-353-585-5	Sequence 5, Appl1
462	6	0.7	211	6	5196194-13	Patent No. 5196194	535	6	0.7	303	1	US-07-797-553-4	Sequence 4, Appl1
463	6	0.7	212	3	US-08-968-747-3	Sequence 3, Appl1	536	6	0.7	306	1	US-08-312-387B-4	Sequence 4, Appl1
464	6	0.7	212	4	US-08-914-999-17	Sequence 17, Appl	537	6	0.7	306	1	US-08-683-426-4	Sequence 4, Appl1
465	6	0.7	213	2	US-08-846-790A-3	Sequence 3, Appl1	538	6	0.7	306	1	US-08-683-458-4	Sequence 4, Appl1

539	6	0.7	306	2	US-08-878-360-4	Sequence 4, Appli	612	6	0.7	389	3	US-09-181-318-3	Sequence 3, Appli
540	6	0.7	306	3	US-08-478-140B-4	Sequence 4, Appli	613	6	0.7	393	1	US-08-423-399B-33	Sequence 33, Appli
541	6	0.7	306	4	US-09-333-412-4	Sequence 4, Appli	614	6	0.7	393	1	US-08-530-950-11	Sequence 11, Appli
542	6	0.7	309	2	US-08-464-517-4	Sequence 4, Appli	615	6	0.7	393	4	US-08-888-429A-11	Sequence 11, Appli
543	6	0.7	309	3	US-08-463-772-4	Sequence 4, Appli	616	6	0.7	393	4	US-09-149-879-11	Sequence 11, Appli
544	6	0.7	317	4	US-08-943-155-6	Sequence 6, Appli	617	6	0.7	396	3	US-08-985-908-24	Sequence 24, Appli
545	6	0.7	318	2	US-08-671-947-2	Sequence 2, Appli	618	6	0.7	397	4	US-09-433-248A-2	Sequence 2, Appli
546	6	0.7	319	5	PCT-US94-04208-2	Sequence 2, Appli	619	6	0.7	397	5	PCT-US96-10602-6	Sequence 6, Appli
547	6	0.7	321	2	US-08-922-146-2	Sequence 2, Appli	620	6	0.7	398	1	US-08-091-519-2	Sequence 2, Appli
548	6	0.7	321	2	US-08-922-146-4	Sequence 2, Appli	621	6	0.7	398	1	US-08-442-043A-2	Sequence 2, Appli
549	6	0.7	321	4	US-09-343-986-2	Sequence 2, Appli	622	6	0.7	398	4	US-09-213-053-3	Sequence 3, Appli
550	6	0.7	321	4	US-09-343-986-4	Sequence 2, Appli	623	6	0.7	398	4	US-09-173-151A-26	Sequence 26, Appli
551	6	0.7	326	1	US-07-603-133B-28	Sequence 28, Appli	624	6	0.7	398	5	PCT-US91-03478-2	Sequence 2, Appli
552	6	0.7	330	1	US-08-238-163-2	Sequence 2, Appli	625	6	0.7	400	2	US-08-878-989-17	Sequence 17, Appli
553	6	0.7	334	2	US-08-484-397A-8	Sequence 8, Appli	626	6	0.7	400	4	US-09-272-796-17	Sequence 17, Appli
554	6	0.7	339	1	US-08-266-451B-22	Sequence 22, Appli	627	6	0.7	402	4	US-09-319-892-4	Sequence 4, Appli
555	6	0.7	339	2	US-08-748-725-22	Sequence 22, Appli	628	6	0.7	402	4	US-09-464-483-4	Sequence 4, Appli
556	6	0.7	342	4	US-09-381-810A-1	Sequence 1, Appli	629	6	0.7	402	4	US-09-414-664-4	Sequence 4, Appli
557	6	0.7	343	2	US-08-282-197C-57	Sequence 57, Appli	630	6	0.7	404	1	US-08-696-770-2	Sequence 2, Appli
558	6	0.7	344	1	US-08-891-234-7	Sequence 7, Appli	631	6	0.7	404	2	US-09-015-557-2	Sequence 2, Appli
559	6	0.7	344	2	US-08-819-539-7	Sequence 7, Appli	632	6	0.7	406	1	US-08-487-823B-5	Sequence 5, Appli
560	6	0.7	344	2	US-09-030-270A-7	Sequence 7, Appli	633	6	0.7	406	2	US-08-997-040-5	Sequence 5, Appli
561	6	0.7	344	4	US-08-984-207-7	Sequence 7, Appli	634	6	0.7	406	2	US-09-203-237-5	Sequence 5, Appli
562	6	0.7	344	4	US-09-013-587-7	Sequence 7, Appli	635	6	0.7	409	4	US-09-572-191-6	Sequence 6, Appli
563	6	0.7	344	5	PCT-US96-08819-7	Sequence 7, Appli	636	6	0.7	410	2	US-08-449-986-2	Sequence 2, Appli
564	6	0.7	346	1	US-08-105-483-217	Sequence 217, App	637	6	0.7	410	2	US-08-756-855-2	Sequence 2, Appli
565	6	0.7	346	1	US-08-709-209-217	Sequence 217, App	638	6	0.7	413	4	US-08-942-572-2	Sequence 2, Appli
566	6	0.7	346	1	US-08-458-101-217	Sequence 217, App	639	6	0.7	420	1	US-08-487-823B-4	Sequence 4, Appli
567	6	0.7	347	3	US-08-749-816-3	Sequence 3, Appli	640	6	0.7	420	1	US-08-997-040-4	Sequence 4, Appli
568	6	0.7	347	4	US-09-144-914-6	Sequence 6, Appli	641	6	0.7	420	2	US-09-203-237-4	Sequence 4, Appli
569	6	0.7	348	1	US-08-366-953A-45	Sequence 45, Appli	642	6	0.7	423	2	US-08-290-731C-10	Sequence 10, Appli
570	6	0.7	348	2	US-08-484-397A-2	Sequence 2, Appli	643	6	0.7	423	2	US-08-290-731C-11	Sequence 11, Appli
571	6	0.7	348	2	US-08-484-397A-3	Sequence 3, Appli	644	6	0.7	425	1	US-08-414-926A-15	Sequence 15, Appli
572	6	0.7	348	2	US-08-484-397A-4	Sequence 4, Appli	645	6	0.7	425	2	US-08-926-922-15	Sequence 15, Appli
573	6	0.7	348	2	US-08-484-397A-5	Sequence 5, Appli	646	6	0.7	425	3	US-09-253-682-15	Sequence 15, Appli
574	6	0.7	348	2	US-08-484-397A-6	Sequence 6, Appli	647	6	0.7	425	4	US-09-527-657-15	Sequence 15, Appli
575	6	0.7	348	2	US-08-484-397A-7	Sequence 7, Appli	648	6	0.7	426	4	US-09-228-986-124	Sequence 124, App
576	6	0.7	348	2	US-08-484-397A-27	Sequence 27, Appli	649	6	0.7	427	4	US-09-182-816-16	Sequence 16, Appli
577	6	0.7	348	2	US-08-484-397A-38	Sequence 38, Appli	650	6	0.7	427	4	US-09-471-528-16	Sequence 16, Appli
578	6	0.7	351	5	PCT-US96-10602-4	Sequence 4, Appli	651	6	0.7	427	4	US-09-634-530-16	Sequence 16, Appli
579	6	0.7	352	1	US-08-482-577B-4	Sequence 4, Appli	652	6	0.7	430	4	US-09-182-816-28	Sequence 28, Appli
580	6	0.7	352	4	US-09-218-176-4	Sequence 4, Appli	653	6	0.7	430	4	US-09-471-528-28	Sequence 28, Appli
581	6	0.7	353	3	US-08-986-485-6	Sequence 6, Appli	654	6	0.7	430	4	US-09-634-530-28	Sequence 28, Appli
582	6	0.7	361	2	US-08-209-521-5	Sequence 5, Appli	655	6	0.7	434	4	US-08-630-915A-22	Sequence 22, Appli
583	6	0.7	361	4	US-08-961-810-1	Sequence 1, Appli	656	6	0.7	439	4	US-08-975-762-60	Sequence 60, Appli
584	6	0.7	361	4	US-08-352-902D-1	Sequence 1, Appli	657	6	0.7	439	4	US-09-295-028-60	Sequence 60, Appli
585	6	0.7	364	1	US-08-318-831-6	Sequence 6, Appli	658	6	0.7	439	4	US-09-106-582-60	Sequence 60, Appli
586	6	0.7	366	3	US-08-746-883-6	Sequence 6, Appli	659	6	0.7	442	3	US-09-032-365A-17	Sequence 17, Appli
587	6	0.7	368	1	US-08-303-238-3	Sequence 3, Appli	660	6	0.7	442	4	US-09-471-528-33	Sequence 33, Appli
588	6	0.7	368	1	US-08-423-399B-35	Sequence 35, Appli	661	6	0.7	442	4	US-09-634-530-33	Sequence 33, Appli
589	6	0.7	368	4	US-08-458-834-3	Sequence 3, Appli	662	6	0.7	451	1	US-08-435-454-4	Sequence 4, Appli
590	6	0.7	368	6	5340934-2	Patent No. 5340934	663	6	0.7	451	1	US-08-652-972A-4	Sequence 4, Appli
591	6	0.7	373	3	US-08-746-883-4	Sequence 4, Appli	664	6	0.7	451	2	US-08-919-145-6	Sequence 6, Appli
592	6	0.7	373	4	US-09-359-161-3	Sequence 3, Appli	665	6	0.7	451	3	US-08-870-126-4	Sequence 4, Appli
593	6	0.7	374	2	US-08-915-107-2	Sequence 2, Appli	666	6	0.7	451	4	US-09-344-889-6	Sequence 6, Appli
594	6	0.7	374	2	US-08-915-107-4	Sequence 4, Appli	667	6	0.7	451	5	PCT-US96-06231A-4	Sequence 4, Appli
595	6	0.7	374	4	US-09-273-613-2	Sequence 2, Appli	668	6	0.7	456	2	US-08-910-731-2	Sequence 2, Appli
596	6	0.7	374	4	US-09-273-613-4	Sequence 4, Appli	669	6	0.7	456	2	US-08-910-731-4	Sequence 4, Appli
597	6	0.7	374	4	US-09-310-867-2	Sequence 2, Appli	670	6	0.7	456	2	US-08-910-731-8	Sequence 8, Appli
598	6	0.7	374	4	US-09-310-867-4	Sequence 4, Appli	671	6	0.7	456	2	US-08-795-395-2	Sequence 2, Appli
599	6	0.7	375	1	US-08-121-714-5	Sequence 5, Appli	672	6	0.7	456	2	US-08-795-395-4	Sequence 4, Appli
600	6	0.7	375	1	US-08-205-719-2	Sequence 2, Appli	673	6	0.7	459	3	US-08-971-782-2	Sequence 2, Appli
601	6	0.7	375	1	US-08-477-108A-5	Sequence 5, Appli	674	6	0.7	459	4	US-09-309-026-2	Sequence 2, Appli
602	6	0.7	375	2	US-08-883-534-1	Sequence 1, Appli	675	6	0.7	464	3	US-08-989-510A-8	Sequence 8, Appli
603	6	0.7	375	2	US-08-477-112-5	Sequence 5, Appli	676	6	0.7	464	4	US-09-182-816-8	Sequence 8, Appli
604	6	0.7	375	3	US-09-204-764-1	Sequence 1, Appli	677	6	0.7	464	4	US-09-182-816-11	Sequence 11, Appli
605	6	0.7	375	3	US-08-746-883-5	Sequence 5, Appli	678	6	0.7	464	4	US-09-471-528-8	Sequence 8, Appli
606	6	0.7	375	4	US-09-572-191-4	Sequence 4, Appli	679	6	0.7	464	4	US-09-471-528-11	Sequence 11, Appli
607	6	0.7	375	5	PCT-US93-08322-5	Sequence 5, Appli	680	6	0.7	464	4	US-09-634-530-8	Sequence 8, Appli
608	6	0.7	380	3	US-08-971-782-4	Sequence 4, Appli	681	6	0.7	464	4	US-09-634-530-11	Sequence 11, Appli
609	6	0.7	380	4	US-09-309-026-4	Sequence 4, Appli	682	6	0.7	464	5	PCT-US96-10521-18	Sequence 18, Appli
610	6	0.7	387	4	US-09-364-230-18	Sequence 18, Appli	683	6	0.7	465	4	US-09-182-816-23	Sequence 23, Appli
611	6	0.7	389	1	US-08-650-275-3	Sequence 3, Appli	684	6	0.7	465	4	US-09-471-528-23	Sequence 23, Appli

685	6	0.7	465	4	US-08-845-258-30	Sequence 30, Appl	758	6	0.7	581	4	US-09-234-393-20	Sequence 20, Appl
686	6	0.7	465	4	US-08-990-571-30	Sequence 30, Appl	759	6	0.7	581	4	US-09-234-393-44	Sequence 44, Appl
687	6	0.7	465	4	US-09-634-530-23	Sequence 23, Appl	760	6	0.7	581	4	US-09-234-393-46	Sequence 46, Appl
688	6	0.7	465	4	US-08-868-435-33	Sequence 33, Appl	761	6	0.7	584	1	US-08-161-290-1	Sequence 1, Appl
689	6	0.7	465	4	US-08-723-142A-30	Sequence 30, Appl	762	6	0.7	584	1	US-08-448-196A-7	Sequence 7, Appl
690	6	0.7	465	4	US-08-744-231-33	Sequence 33, Appl	763	6	0.7	584	1	US-08-450-755-1	Sequence 1, Appl
691	6	0.7	467	4	US-08-495-484-12	Sequence 12, Appl	764	6	0.7	585	1	US-08-117-907-2	Sequence 2, Appl
692	6	0.7	471	2	US-08-477-451-20	Sequence 20, Appl	765	6	0.7	585	1	US-08-485-718-11	Sequence 11, Appl
693	6	0.7	476	4	US-08-565-655-6	Sequence 6, Appl	766	6	0.7	585	1	US-08-485-718-13	Sequence 13, Appl
694	6	0.7	478	4	US-09-069-023-4	Sequence 4, Appl	767	6	0.7	585	2	US-08-484-530-57	Sequence 57, Appl
695	6	0.7	479	2	US-08-807-200-12	Sequence 12, Appl	768	6	0.7	585	2	US-08-484-530-59	Sequence 59, Appl
696	6	0.7	479	3	US-08-852-782-3	Sequence 3, Appl	769	6	0.7	585	2	US-08-494-624-2	Sequence 2, Appl
697	6	0.7	479	4	US-09-001-777-12	Sequence 12, Appl	770	6	0.7	585	2	US-08-827-618A-57	Sequence 57, Appl
698	6	0.7	479	5	PCT-US96-10521-7	Sequence 7, Appl	771	6	0.7	585	2	US-08-827-618A-59	Sequence 59, Appl
699	6	0.7	484	1	US-08-361-611-4	Sequence 4, Appl	772	6	0.7	585	3	US-08-483-952A-57	Sequence 57, Appl
700	6	0.7	484	2	US-08-565-655-4	Sequence 4, Appl	773	6	0.7	585	3	US-08-483-952A-59	Sequence 59, Appl
701	6	0.7	484	3	US-08-946-967-4	Sequence 4, Appl	774	6	0.7	585	3	US-08-453-040-2	Sequence 2, Appl
702	6	0.7	486	3	US-08-904-452-2	Sequence 2, Appl	775	6	0.7	585	3	US-09-043-930-3	Sequence 3, Appl
703	6	0.7	489	2	US-08-663-566A-6	Sequence 6, Appl	776	6	0.7	585	3	US-09-043-930-4	Sequence 4, Appl
704	6	0.7	489	2	US-08-023-610-6	Sequence 6, Appl	777	6	0.7	585	3	US-09-043-930-5	Sequence 5, Appl
705	6	0.7	489	2	US-08-288-065A-6	Sequence 6, Appl	778	6	0.7	585	3	US-09-043-930-6	Sequence 6, Appl
706	6	0.7	489	2	US-08-362-240A-6	Sequence 6, Appl	779	6	0.7	585	3	US-09-043-930-7	Sequence 7, Appl
707	6	0.7	489	4	US-08-804-372A-4	Sequence 4, Appl	780	6	0.7	585	3	US-09-043-930-8	Sequence 8, Appl
708	6	0.7	489	5	PCT-US95-10245-6	Sequence 6, Appl	781	6	0.7	585	6	5475086-6	Patent No. 5475086
709	6	0.7	490	1	US-08-361-611-2	Sequence 2, Appl	782	6	0.7	588	1	US-08-391-615-5	Sequence 5, Appl
710	6	0.7	490	1	US-08-565-655-2	Sequence 2, Appl	783	6	0.7	603	1	US-08-190-802A-50	Sequence 50, Appl
711	6	0.7	490	2	US-08-946-967-2	Sequence 2, Appl	784	6	0.7	603	4	US-08-477-345-50	Sequence 50, Appl
712	6	0.7	491	4	US-08-812-824-1	Sequence 1, Appl	785	6	0.7	603	4	US-08-473-089-50	Sequence 50, Appl
713	6	0.7	492	4	US-09-183-953-2	Sequence 2, Appl	786	6	0.7	605	1	US-08-190-802A-49	Sequence 49, Appl
714	6	0.7	495	3	US-08-962-859A-2	Sequence 2, Appl	787	6	0.7	605	4	US-09-063-950-5	Sequence 5, Appl
715	6	0.7	496	1	US-08-665-220-4	Sequence 4, Appl	788	6	0.7	605	4	US-08-477-346-49	Sequence 49, Appl
716	6	0.7	496	4	US-09-291-692-4	Sequence 4, Appl	789	6	0.7	605	4	US-08-473-089-49	Sequence 49, Appl
717	6	0.7	497	4	US-09-058-947A-4	Sequence 4, Appl	790	6	0.7	609	1	US-07-798-776-2	Sequence 2, Appl
718	6	0.7	500	2	US-09-031-392-7	Sequence 7, Appl	791	6	0.7	609	3	US-08-251-288A-2	Sequence 2, Appl
719	6	0.7	500	4	US-09-299-549-7	Sequence 7, Appl	792	6	0.7	609	3	US-09-298-819A-2	Sequence 2, Appl
720	6	0.7	500	4	US-08-868-373-12	Sequence 12, Appl	793	6	0.7	615	4	US-08-676-444-44	Sequence 44, Appl
721	6	0.7	500	4	US-09-610-417-7	Sequence 7, Appl	794	6	0.7	615	1	US-08-472-934-4	Sequence 4, Appl
722	6	0.7	511	4	US-09-201-641-2	Sequence 2, Appl	795	6	0.7	619	1	US-08-472-934-12	Sequence 12, Appl
723	6	0.7	523	2	US-08-473-553A-3	Sequence 3, Appl	796	6	0.7	619	2	US-08-323-460A-4	Sequence 4, Appl
724	6	0.7	527	2	US-08-365-486A-26	Sequence 26, Appl	797	6	0.7	619	2	US-08-461-146C-4	Sequence 4, Appl
725	6	0.7	527	4	US-08-880-342-26	Sequence 26, Appl	798	6	0.7	619	2	US-08-461-146C-12	Sequence 12, Appl
726	6	0.7	529	4	US-09-464-483-2	Sequence 2, Appl	799	6	0.7	619	3	US-08-461-145C-4	Sequence 4, Appl
727	6	0.7	529	4	US-09-414-664-2	Sequence 2, Appl	800	6	0.7	619	3	US-08-461-145C-12	Sequence 12, Appl
728	6	0.7	530	4	US-09-069-023-3	Sequence 3, Appl	801	6	0.7	619	4	US-09-423-890-4	Sequence 4, Appl
729	6	0.7	531	4	US-09-069-023-1	Sequence 1, Appl	802	6	0.7	619	4	US-09-423-890-10	Sequence 10, Appl
730	6	0.7	533	2	US-08-770-544-4	Sequence 4, Appl	803	6	0.7	619	4	US-08-628-823-6	Sequence 6, Appl
731	6	0.7	534	1	US-08-577-184-2	Sequence 2, Appl	804	6	0.7	619	4	US-08-628-823-8	Sequence 8, Appl
732	6	0.7	539	2	US-08-735-041A-2	Sequence 2, Appl	805	6	0.7	621	1	US-08-295-814E-6	Sequence 6, Appl
733	6	0.7	539	3	US-09-190-476B-2	Sequence 2, Appl	806	6	0.7	621	4	US-09-343-361-6	Sequence 6, Appl
734	6	0.7	539	3	US-09-190-889A-2	Sequence 2, Appl	807	6	0.7	621	5	PCT-US93-01959-6	Sequence 6, Appl
735	6	0.7	539	4	US-09-190-938B-2	Sequence 2, Appl	808	6	0.7	630	4	US-09-228-986-71	Sequence 71, Appl
736	6	0.7	539	5	PCT-US95-09261-2	Sequence 2, Appl	809	6	0.7	640	3	US-09-026-343-2	Sequence 2, Appl
737	6	0.7	540	3	US-09-019-943-1	Sequence 1, Appl	810	6	0.7	641	4	US-09-422-869-26	Sequence 26, Appl
738	6	0.7	540	4	US-09-099-041A-2	Sequence 2, Appl	811	6	0.7	643	2	US-08-245-511-47	Sequence 47, Appl
739	6	0.7	540	4	US-09-069-023-27	Sequence 27, Appl	812	6	0.7	643	2	US-08-600-993A-47	Sequence 47, Appl
740	6	0.7	542	1	US-08-321-587-2	Sequence 2, Appl	813	6	0.7	652	1	US-08-318-831-8	Sequence 8, Appl
741	6	0.7	544	2	US-08-587-680A-25	Patent No. 5256558	814	6	0.7	655	2	US-08-272-255-18	Sequence 18, Appl
742	6	0.7	548	6	5256558-5	Sequence 25, Appl	815	6	0.7	655	4	US-09-228-986-70	Sequence 70, Appl
743	6	0.7	552	4	US-08-796-899-27	Sequence 27, Appl	816	6	0.7	655	5	PCT-US95-08565-18	Sequence 18, Appl
744	6	0.7	554	4	US-08-591-468-7	Sequence 7, Appl	817	6	0.7	662	4	US-08-779-814-5	Sequence 5, Appl
745	6	0.7	554	5	PCT-US94-06430-7	Sequence 7, Appl	818	6	0.7	673	4	US-09-063-950-2	Sequence 2, Appl
746	6	0.7	555	2	US-08-982-232-7	Sequence 7, Appl	819	6	0.7	681	4	US-08-760-615-4	Sequence 4, Appl
747	6	0.7	555	2	US-08-982-232-14	Sequence 14, Appl	820	6	0.7	693	1	US-08-553-279-2	Sequence 2, Appl
748	6	0.7	560	2	US-08-095-728B-6	Sequence 6, Appl	821	6	0.7	707	4	US-09-228-986-80	Sequence 80, Appl
749	6	0.7	560	3	US-08-592-500-2	Sequence 2, Appl	822	6	0.7	708	4	US-09-131-648-2	Sequence 2, Appl
750	6	0.7	560	3	US-08-193-006-2	Sequence 2, Appl	823	6	0.7	712	1	US-08-587-889-2	Sequence 2, Appl
751	6	0.7	560	3	US-08-983-045-4	Sequence 4, Appl	824	6	0.7	712	2	US-08-587-889-2	Sequence 2, Appl
752	6	0.7	560	4	US-09-063-950-4	Sequence 4, Appl	825	6	0.7	712	2	US-09-307-185-5	Sequence 5, Appl
753	6	0.7	560	5	PCT-US92-02320A-6	Sequence 6, Appl	826	6	0.7	712	5	PCT-US96-09193-2	Sequence 2, Appl
754	6	0.7	568	4	US-09-07644A-2	Sequence 2, Appl	827	6	0.7	727	3	US-08-482-677-11	Sequence 11, Appl
755	6	0.7	568	4	US-09-238-373-2	Sequence 2, Appl	828	6	0.7	727	3	US-08-650-599A-4	Sequence 4, Appl
756	6	0.7	572	4	US-09-237-111-2	Sequence 2, Appl	829	6	0.7	727	4	US-09-490-517-4	Sequence 4, Appl
757	6	0.7	580	4	US-09-234-393-48	Sequence 48, Appl	830	6	0.7	729	2	US-08-677-298-2	Sequence 2, Appl

831	6	0.7	732	4	US-08-914-999-8	Sequence 8, Appli	904	6	0.7	910	4	US-08-460-269C-2	Sequence 2, Appli
832	6	0.7	735	3	US-09-191-647-9	Sequence 9, Appli	905	6	0.7	910	4	US-09-228-986-72	Sequence 72, Appli
833	6	0.7	735	4	US-09-540-245A-9	Sequence 9, Appli	906	6	0.7	912	4	US-08-943-768-2	Sequence 2, Appli
834	6	0.7	735	4	US-09-540-153-9	Sequence 9, Appli	907	6	0.7	922	4	US-09-141-206-6	Sequence 6, Appli
835	6	0.7	740	1	US-08-016-863-16	Sequence 16, Appli	908	6	0.7	926	3	US-08-755-587-187	Sequence 187, Appli
836	6	0.7	740	3	US-08-276-968A-16	Sequence 16, Appli	909	6	0.7	932	3	US-08-968-752B-6	Sequence 6, Appli
837	6	0.7	740	3	US-08-276-968A-18	Sequence 18, Appli	910	6	0.7	933	4	US-09-107-149-2	Sequence 2, Appli
838	6	0.7	741	3	US-08-276-968A-20	Sequence 20, Appli	911	6	0.7	933	4	US-09-107-149-2	Sequence 17, Appli
839	6	0.7	746	5	PCT-US95-10509-2	Sequence 2, Appli	912	6	0.7	942	2	US-08-884-681-4	Sequence 4, Appli
840	6	0.7	746	5	PCT-US95-10509-2	Sequence 2, Appli	913	6	0.7	942	2	US-08-884-681-4	Sequence 4, Appli
841	6	0.7	747	3	US-09-035-648-18	Sequence 18, Appli	914	6	0.7	942	2	US-08-258-643-4	Sequence 73, Appli
842	6	0.7	747	4	US-09-001-951-18	Sequence 18, Appli	915	6	0.7	947	4	US-09-228-986-73	Sequence 3, Appli
843	6	0.7	748	4	US-08-392-459-24	Sequence 24, Appli	916	6	0.7	968	4	US-09-180-439-3	Sequence 4, Appli
844	6	0.7	748	4	US-08-392-459-28	Sequence 28, Appli	917	6	0.7	968	4	US-09-180-439-3	Sequence 4, Appli
845	6	0.7	748	5	PCT-US91-08525-24	Sequence 24, Appli	918	6	0.7	968	4	US-09-228-986-76	Sequence 76, Appli
846	6	0.7	748	5	PCT-US91-08525-28	Sequence 28, Appli	919	6	0.7	969	2	US-08-284-941-2	Sequence 2, Appli
847	6	0.7	748	5	PCT-US91-08525-34	Sequence 34, Appli	920	6	0.7	969	2	US-08-447-642-2	Sequence 2, Appli
848	6	0.7	748	5	PCT-US93-04384-4	Sequence 4, Appli	921	6	0.7	969	2	US-09-236-503-2	Sequence 2, Appli
849	6	0.7	748	5	PCT-US93-04384-6	Sequence 6, Appli	922	6	0.7	969	5	PCT-US93-02147A-2	Sequence 2, Appli
850	6	0.7	758	1	US-08-289-112-2	Sequence 2, Appli	923	6	0.7	970	1	US-08-375-709-7	Sequence 7, Appli
851	6	0.7	764	4	US-07-757-342D-5	Sequence 5, Appli	924	6	0.7	970	1	US-08-752-929-7	Sequence 7, Appli
852	6	0.7	764	4	US-07-741-453A-59	Sequence 59, Appli	925	6	0.7	970	4	US-09-090-793-5	Sequence 5, Appli
853	6	0.7	764	4	US-07-741-453A-61	Sequence 61, Appli	926	6	0.7	971	2	US-08-724-354D-22	Sequence 22, Appli
854	6	0.7	774	4	US-09-009-191-4	Sequence 4, Appli	927	6	0.7	971	3	US-09-270-984A-22	Sequence 22, Appli
855	6	0.7	790	2	US-08-359-705B-9	Sequence 9, Appli	928	6	0.7	971	3	US-09-177-431-8	Sequence 8, Appli
856	6	0.7	790	2	US-08-286-846A-9	Sequence 9, Appli	929	6	0.7	973	3	US-08-904-452-4	Sequence 4, Appli
857	6	0.7	790	2	US-08-457-880A-9	Sequence 9, Appli	930	6	0.7	980	2	US-08-473-553A-6	Sequence 6, Appli
858	6	0.7	790	3	US-08-444-622A-9	Sequence 9, Appli	931	6	0.7	985	2	US-08-473-553A-2	Sequence 2, Appli
859	6	0.7	790	3	US-08-942-562-9	Sequence 9, Appli	932	6	0.7	985	3	US-08-938-291A-7	Sequence 7, Appli
860	6	0.7	790	4	US-08-156-923-9	Sequence 9, Appli	933	6	0.7	995	5	PCT-US93-04910-14	Sequence 14, Appli
861	6	0.7	791	4	US-08-861-745B-1	Sequence 1, Appli	934	6	0.7	997	2	US-08-387-942C-4	Sequence 4, Appli
862	6	0.7	797	2	US-08-095-728B-2	Sequence 2, Appli	935	6	0.7	1000	4	US-09-193-562D-30	Sequence 30, Appli
863	6	0.7	797	5	PCT-US92-02320A-2	Sequence 2, Appli	936	6	0.7	1013	4	US-09-415-522-8	Sequence 8, Appli
864	6	0.7	799	4	US-09-180-439-6	Sequence 6, Appli	937	6	0.7	1013	4	US-08-860-886-2	Sequence 2, Appli
865	6	0.7	800	6	5183745-3	Patent No. 5183745	938	6	0.7	1014	4	US-09-078-347A-3	Sequence 3, Appli
866	6	0.7	814	1	US-08-286-305A-3	Sequence 3, Appli	939	6	0.7	1016	4	US-09-180-439-8	Sequence 8, Appli
867	6	0.7	814	2	US-08-441-104A-3	Sequence 3, Appli	940	6	0.7	1023	1	US-08-198-446B-4	Sequence 4, Appli
868	6	0.7	814	2	US-08-440-816A-3	Sequence 3, Appli	941	6	0.7	1023	2	US-08-870-693-4	Sequence 4, Appli
869	6	0.7	814	4	US-09-417-381A-3	Sequence 3, Appli	942	6	0.7	1036	2	US-08-720-484A-5	Sequence 5, Appli
870	6	0.7	837	3	US-09-012-710-12	Sequence 12, Appli	943	6	0.7	1036	4	US-08-953-823A-5	Sequence 5, Appli
871	6	0.7	837	4	US-09-556-273-12	Sequence 12, Appli	944	6	0.7	1053	3	US-08-613-009A-8	Sequence 8, Appli
872	6	0.7	839	4	US-09-197-636-2	Sequence 2, Appli	945	6	0.7	1066	2	US-08-308-818-1	Sequence 1, Appli
873	6	0.7	839	4	US-09-197-636-8	Sequence 4, Appli	946	6	0.7	1070	3	US-08-613-009A-11	Sequence 11, Appli
874	6	0.7	839	4	US-09-197-636-8	Sequence 4, Appli	947	6	0.7	1073	4	US-09-541-782-6	Sequence 6, Appli
875	6	0.7	839	4	US-09-235-451-34	Sequence 34, Appli	948	6	0.7	1074	3	US-08-613-009A-7	Sequence 7, Appli
876	6	0.7	840	4	US-08-974-549A-190	Sequence 190, App	949	6	0.7	1076	2	US-08-867-941-19	Sequence 19, Appli
877	6	0.7	843	4	US-09-235-451-25	Sequence 25, Appli	950	6	0.7	1076	4	US-09-074-658-19	Sequence 19, Appli
878	6	0.7	847	1	US-08-276-099A-2	Sequence 2, Appli	951	6	0.7	1091	3	US-08-986-485-5	Sequence 5, Appli
879	6	0.7	847	1	US-08-781-890-2	Sequence 2, Appli	952	6	0.7	1101	3	US-08-986-485-2	Sequence 2, Appli
880	6	0.7	847	4	US-09-087-465-12	Sequence 12, Appli	953	6	0.7	1112	4	US-09-353-585-2	Sequence 2, Appli
881	6	0.7	860	1	US-08-117-362-3	Sequence 3, Appli	954	6	0.7	1112	4	US-09-353-585-3	Sequence 3, Appli
882	6	0.7	860	1	US-08-486-924-3	Sequence 3, Appli	955	6	0.7	1130	2	US-08-519-547A-6	Sequence 6, Appli
883	6	0.7	863	1	US-07-923-976-8	Sequence 8, Appli	956	6	0.7	1147	1	US-08-144-121-3	Sequence 3, Appli
884	6	0.7	863	2	US-08-380-182-19	Sequence 19, Appli	957	6	0.7	1147	2	US-08-735-893-3	Sequence 3, Appli
885	6	0.7	863	2	US-08-380-182-20	Sequence 20, Appli	958	6	0.7	1165	1	US-08-144-121-2	Sequence 2, Appli
886	6	0.7	872	4	US-08-844-057-2	Sequence 2, Appli	959	6	0.7	1165	2	US-08-735-893-2	Sequence 2, Appli
887	6	0.7	872	4	US-08-387-942C-5	Sequence 5, Appli	960	6	0.7	1184	4	US-09-541-782-2	Sequence 2, Appli
888	6	0.7	872	3	US-08-851-843A-8	Sequence 8, Appli	961	6	0.7	1220	3	US-08-930-996A-2	Sequence 2, Appli
889	6	0.7	872	3	US-08-851-843A-54	Sequence 54, Appli	962	6	0.7	1240	3	US-08-930-996A-4	Sequence 4, Appli
890	6	0.7	872	4	US-08-974-549A-221	Sequence 221, App	963	6	0.7	1262	4	US-09-357-251-33	Sequence 33, Appli
891	6	0.7	872	4	US-08-854-050-8	Sequence 8, Appli	964	6	0.7	1263	4	US-09-446-504-6	Sequence 6, Appli
892	6	0.7	872	4	US-08-854-050-54	Sequence 54, Appli	965	6	0.7	1263	4	US-09-712-266-6	Sequence 6, Appli
893	6	0.7	872	4	US-09-430-323-8	Sequence 8, Appli	966	6	0.7	1266	1	US-08-468-557-4	Sequence 4, Appli
894	6	0.7	872	4	US-09-430-323-54	Sequence 54, Appli	967	6	0.7	1266	4	US-09-357-251-32	Sequence 32, Appli
895	6	0.7	897	1	US-08-095-737-4	Sequence 4, Appli	968	6	0.7	1287	1	US-08-200-232-2	Sequence 2, Appli
896	6	0.7	897	1	US-08-480-145-4	Sequence 4, Appli	969	6	0.7	1287	5	PCT-US95-02219-2	Sequence 2, Appli
897	6	0.7	897	2	US-08-477-389-4	Sequence 4, Appli	970	6	0.7	1287	5	PCT-US95-02219A-2	Sequence 2, Appli
898	6	0.7	900	2	US-08-630-822A-62	Sequence 62, Appli	971	6	0.7	1297	2	US-08-290-731C-2	Sequence 2, Appli
899	6	0.7	900	2	US-09-005-069-62	Sequence 62, Appli	972	6	0.7	1319	2	US-08-290-731C-2	Sequence 2, Appli
900	6	0.7	901	2	US-08-884-681-5	Sequence 5, Appli	973	6	0.7	1332	4	US-08-971-244-2	Sequence 2, Appli
901	6	0.7	901	4	US-09-258-643-5	Sequence 5, Appli	974	6	0.7	1332	4	US-09-286-891-2	Sequence 2, Appli
902	6	0.7	907	3	US-08-783-774-2	Sequence 2, Appli	975	6	0.7	1333	3	US-09-356-952-2	Sequence 2, Appli
903	6	0.7	907	5	PCT-US95-04611A-19	Sequence 19, Appli	976	6	0.7	1336	2	US-08-290-731C-6	Sequence 6, Appli
										1377	2	US-08-308-818-4	Sequence 4, Appli



; Patent No. 5705609  
; GENERAL INFORMATION:  
; APPLICANT: RUOSLAHTI, ERKKI I.  
; APPLICANT: PIERSCHBACHER, MICHAEL D.  
; APPLICANT: CARDENAS, JOSE  
; APPLICANT: CRAIG, WILLIAM  
; APPLICANT: MULLEN, DANIEL G.  
; TITLE OF INVENTION: DECORIN FRAGMENTS AND METHODS OF  
; TITLE OF INVENTION: INHIBITING CELL REGULATORY FACTORS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/442,063A  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/865,652  
; FILING DATE: 03-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 1454  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-535-9001  
; TELEFAX: 619-535-8949  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 307 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-442-063A-48

Query Match 1.1%; Score 9; DB 1; Length 307;  
Best Local Similarity 100.0%; Pred. No. 2.5; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 113 IPQGLPPSL 121  
| | | | |  
DB 187 IPQGLPPSL 195

RESULT 4  
US-07-613-083B-1  
; Sequence 1, Application US/07613083B  
; Patent No. 5340727  
; GENERAL INFORMATION:  
; APPLICANT: Ruggeri, Zaverio M.  
; APPLICANT: Ware, Jerry, inventors  
; APPLICANT: on behalf of Scripps Clinic and Research  
; APPLICANT: Foundation  
; TITLE OF INVENTION: GPTb' Fragments and Recombinant  
; TITLE OF INVENTION: DNA Expression Vectors  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scripps Clinic and Research  
; ADDRESSEE: Foundation  
; STREET: 10666 No. 5340727th Torrey Pines Road  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA

; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb  
; COMPUTER: AST Bravo IBM PC comp. (386SX)  
; OPERATING SYSTEM: MS DOS version 3.2  
; SOFTWARE: WordPerfect 5.1 conv. to ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/613,083B  
; FILING DATE: 19911114  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: This appl. is a c-i-p of  
; APPLICATION NUMBER: U.S. 07/470,674  
; FILING DATE: 04-Jan-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barron, Alexis  
; REGISTRATION NUMBER: 22,702  
; REFERENCE/DOCKET NUMBER: P16,569-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 923-4466  
; TELEFAX: (215) 923-2189  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 320  
; TYPE: AMINO ACID  
; STRANDEDNESS: No. 5340727 applicable  
; TOPOLOGY: Linear  
; US-07-613-083B-1

Query Match 1.1%; Score 9; DB 1; Length 320;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNOL 682  
| | | | |  
DB 81 TLDLSHNOL 89

RESULT 5  
US-08-442-063A-27  
; Sequence 27, Application US/08442063A  
; Patent No. 5705609  
; GENERAL INFORMATION:  
; APPLICANT: RUOSLAHTI, ERKKI I.  
; APPLICANT: PIERSCHBACHER, MICHAEL D.  
; APPLICANT: CARDENAS, JOSE  
; APPLICANT: CRAIG, WILLIAM  
; APPLICANT: MULLEN, DANIEL G.  
; TITLE OF INVENTION: DECORIN FRAGMENTS AND METHODS OF  
; TITLE OF INVENTION: INHIBITING CELL REGULATORY FACTORS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/442,063A  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/865,652  
; FILING DATE: 03-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN A.  
; REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LA 1454  
TELEPHONE: 619-535-9001  
TELEFAX: 619-535-8949  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 333 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-442-063A-27

Query Match 1.1%; Score 9; DB 1; Length 333;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121  
Db 187 IPQGLPPSL 195

RESULT 6  
US-08-272-919-2  
Sequence 2, Application US/08272919  
Patent No. 5567807  
GENERAL INFORMATION:  
APPLICANT: Craig, William S.  
APPLICANT: Harper, John R.  
APPLICANT: Hernandez, Sam D.  
APPLICANT: Kostel, Paul J.  
APPLICANT: Parker, Jonathan R.  
APPLICANT: Vedvick, Thomas S.  
TITLE OF INVENTION: Processes for the Purification of Human  
TITLE OF INVENTION: Recombinant Decorin and the Detection of Guanidinium Ions  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESS: Campbell and Flores  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/272,919  
FILING DATE: 08-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 1040  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-272-919-2

Query Match 1.1%; Score 9; DB 1; Length 342;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121

Db 199 IPQGLPPSL 207

RESULT 7  
US-08-619-916-2  
Sequence 2, Application US/08619916  
Patent No. 5763276  
GENERAL INFORMATION:  
APPLICANT: Craig, William S.  
APPLICANT: Harper, John R.  
APPLICANT: Hernandez, Sam D.  
APPLICANT: Kostel, Paul J.  
APPLICANT: Parker, Jonathan R.  
APPLICANT: Vedvick, Thomas S.  
TITLE OF INVENTION: Processes for the Purification of Human  
TITLE OF INVENTION: Recombinant Decorin and the Detection of Guanidinium Ions  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESS: Campbell and Flores  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/619,916  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/272,919  
FILING DATE: 08-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 1040  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-619-916-2

Query Match 1.1%; Score 9; DB 1; Length 342;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPQGLPPSL 121  
Db 199 IPQGLPPSL 207

RESULT 8  
PCT-US95-08542-2  
Sequence 2, Application PC/TUS9508542  
GENERAL INFORMATION:  
APPLICANT: La Jolla Cancer Research Foundation  
TITLE OF INVENTION: Processes for the Purification of  
TITLE OF INVENTION: Human  
TITLE OF INVENTION: Recombinant Decorin and the Detection of Guanidinium  
TITLE OF INVENTION: Ions  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:



ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08542  
FILING DATE: 07-JUL-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Susan M.  
REGISTRATION NUMBER: 36,405  
REFERENCE/DOCKET NUMBER: FP-LA 1751  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-08542-2

Query Match 1.1%; Score 9; DB 5; Length 342;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IPQGLPPSL 121  
|||||  
Db 199 IPQGLPPSL 207

RESULT 9  
US-08-303-238-4  
Sequence 4, Application US/08303238  
Patent No. 5654270  
GENERAL INFORMATION:  
APPLICANT: RUOSLAHTI, ERKKI I.  
APPLICANT: LONGAKER, MICHAEL T.  
APPLICANT: WHITBY, DAVID J.  
APPLICANT: HARPER, JOHN R.  
APPLICANT: PIERSCHBACHER, MICHAEL D.  
TITLE OF INVENTION: INHIBITORS OF CELL REGULATORY FACTORS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
CITY: SAN DIEGO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/303,238  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,931  
FILING DATE: 17-NOV-1992

ATTORNEY/AGENT INFORMATION:  
NAME: KONSKI, ANTOINETTE F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: P-LA 9453  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-535-9001  
TELEFAX: 619-535-8949  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-303-238-4

Query Match 1.1%; Score 9; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 IPQGLPPSL 121  
|||||  
Db 215 IPQGLPPSL 223

RESULT 10  
US-08-458-834-4  
Sequence 4, Application US/08458834  
Patent No. 6277812  
GENERAL INFORMATION:  
APPLICANT: RUOSLAHTI, ERKKI I.  
APPLICANT: LONGAKER, MICHAEL T.  
APPLICANT: WHITBY, DAVID J.  
APPLICANT: HARPER, JOHN R.  
APPLICANT: PIERSCHBACHER, MICHAEL D.  
APPLICANT: BORDER, WAYNE A.  
TITLE OF INVENTION: INHIBITORS OF CELL REGULATORY FACTORS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
CITY: SAN DIEGO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,834  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/978,931  
FILING DATE: 17-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: KONSKI, ANTOINETTE F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: P-LA 9453  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-535-9001  
TELEFAX: 619-535-8949  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-458-834-4

Query Match 1.1%; Score 9; DB 4; Length 359;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 IPOGLPPSL 121  
| | | | | | | |  
DB 215 IPOGLPPSL 223

RESULT 11  
US-07-821-717B-6  
; Sequence 6, Application US/07821717B  
; Patent No. 5298239  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; NUMBER OF INVENTIONS: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/07/821,717B  
; FILING DATE: 15-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timain, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PUBLICATION INFORMATION:  
; AUTHORS: Lopez, Jose A.  
; AUTHORS: Chung, Dominic W.  
; AUTHORS: Fujikawa, Kazuo  
; AUTHORS: Hagen, Frederick S.  
; AUTHORS: Papayannopoulou, Thalia  
; AUTHORS: Roth, Gerald J.  
; TITLE: Cloning of the alpha chain of human  
; TITLE: platelet glycoprotein Ib: A transmembrane protein with homology  
; TITLE: to leucine-rich alpha-2-glycoprotein  
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
; VOLUME: 84  
; PAGES: 5615-5619  
; DATE: AUG-1987  
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
; PUBLICATION INFORMATION:  
; AUTHORS: Zimmerman, Theodore S.  
; AUTHORS: Ruggeri, Zaverio M.  
; AUTHORS: Houghten, Richard A.  
; AUTHORS: Vincete, Vincete  
; AUTHORS: Mohri, Hiroshi

; TITLE: proteolytic fragments and synthetic  
; TITLE: peptides that block the binding of von Willebrand factor to the  
; TITLE: platelet membrane glycoprotein Ib  
; DOCUMENT NUMBER: EP 0 317 278 A2  
; FILING DATE: 16-NOV-1988  
; PUBLICATION DATE: 24-MAY-1989  
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-07-821-717B-6

Query Match 1.1%; Score 9; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 674 TLDLSHNL 682  
| | | | | | | |  
DB 81 TLDLSHNL 89

RESULT 12  
US-08-119-262B-6  
; Sequence 6, Application US/08119262B  
; Patent No. 5492809  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/119,262B  
; FILING DATE: 09-SEP-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/821,717  
; FILING DATE: 15-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timain, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PUBLICATION INFORMATION:  
; AUTHORS: Lopez, Jose A.  
; AUTHORS: Chung, Dominic W.  
; AUTHORS: Fujikawa, Kazuo  
; AUTHORS: Hagen, Frederick S.  
; AUTHORS: Papayannopoulou, Thalia  
; AUTHORS: Roth, Gerald J.  
; TITLE: Cloning of the alpha chain of human  
; TITLE: glycoprotein Ib: A transmembrane protein  
; with homol

;; TITLE: leucine-rich alpha-2-glycoprotein  
;; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
;; VOLUME: 84  
;; DATE: 5615-5619  
;; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
;; PUBLICATION INFORMATION:  
;; AUTHORS: Zimmerman, Theodore S.  
;; AUTHORS: Ruggeri, Zaverio M.  
;; AUTHORS: Houghten, Richard A.  
;; AUTHORS: Vincete, Vincete  
;; AUTHORS: Mohri, Hiroshi  
;; TITLE: Proteolytic fragments and synthetic peptides  
;; TITLE: that block the binding of von Willebrand  
;; TITLE: membrane glycoprotein Ib  
;; DOCUMENT NUMBER: EP 0 317 278 A2  
;; FILING DATE: 16-NOV-1988  
;; PUBLICATION DATE: 24-MAY-1989  
;; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
;; US-08-119-262B-6

Query Match 1.1%; Score 9; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 674 TLDLSHNL 682  
Db 81 TLDLSHNL 89

RESULT 13  
US-08-135-929A-11  
; Sequence 11, Application US/08135929A  
; Patent No. 5593959  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; APPLICANT: Pincus, Matthew R.  
; TITLE OF INVENTION: Mutations in the Gene Encoding the Alpha  
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/135,929A  
; FILING DATE: 14-OCT-1993  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/23  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; TELEX: 978450  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

;; MOLECULE TYPE: protein  
US-08-135-929A-11

Query Match 1.1%; Score 9; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 674 TLDLSHNL 682  
Db 81 TLDLSHNL 89

RESULT 14  
US-08-234-265A-11  
; Sequence 11, Application US/08234265A  
; Patent No. 5624817  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; APPLICANT: Pincus, Matthew R.  
; TITLE OF INVENTION: Mutations in the Gene Encoding the  
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/234,265A  
; FILING DATE: 28-APR-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/24  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; TELEX: 978450  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-234-265A-11

Query Match 1.1%; Score 9; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 674 TLDLSHNL 682  
Db 81 TLDLSHNL 89

RESULT 15  
US-09-353-585-8  
; Sequence 8, Application US/09353585  
; Patent No. 6287865  
; GENERAL INFORMATION:

```
;
; APPLICANT: Dixon, Mark S
; Jones, David A
; Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses
; thereof
;
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6287865th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22201-4714
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/353,585
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q
; 1/68
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/930,277
; FILING DATE: 27-OCT-1997
; APPLICATION NUMBER: PCT/GB96/00785
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: GB 9506658.5
; FILING DATE: 31-MAR-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-69
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 65 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: Cf9
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
;
; US-09-353-585-8
;
; Query Match 1.0%; Score 8; DB 4; Length 65;
; Best Local Similarity 100.0%; Pred. NO. 5.5;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 723 LDLSNNKI 730
; Db 1 LDLSNNKI 8
;
; RESULT 16
; US-09-353-585-6
; Sequence 6, Application US/09353585
; Patent No. 6287865
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; Jones, David A
; Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses
; thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
```

```
;
; ADDRESSER: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6287865th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22201-4714
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/353,585
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q
; 1/68
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/930,277
; FILING DATE: 27-OCT-1997
; APPLICATION NUMBER: PCT/GB96/00785
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: GB 9506658.5
; FILING DATE: 31-MAR-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-69
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: Cf9
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
;
; US-09-353-585-6
;
; Query Match 1.0%; Score 8; DB 4; Length 268;
; Best Local Similarity 100.0%; Pred. No. 19;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 723 LDLSNNKI 730
; Db 126 LDLSNNKI 133
;
; RESULT 17
; US-09-171-461-23
; Sequence 23, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbaumer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.180000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 23  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: CELO Virus  
; FEATURE:  
; OTHER INFORMATION: Position: 41002..41853 /note=ORF10  
US-09-171-461-23

Query Match 1.0%; Score 8; DB 4; Length 283;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 TNLTLTN 46  
Db 222 TNLTLTN 229

RESULT 18  
US-08-945-983-2  
; Sequence 2, Application US/08945983  
; Patent No. 622527  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Colwyn M  
; APPLICANT: Balint-Kurti, Peter J  
; APPLICANT: Jones, David A  
; APPLICANT: Jones, Jonathan DG  
; TITLE OF INVENTION: Plant pathogen resistance genes and uses  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye PC  
; STREET: 8th Floor, 1100 No. 622527th Glebe Road  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/945,983  
; FILING DATE: 12-NOV-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB96/01155  
; FILING DATE: 13-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9509575.8  
; FILING DATE: 11-MAY-1995  
; TELECOMMUNICATION INFORMATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ms Mary J Wilson  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-27  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 806 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: lycopersicon hirsutum  
; STRAIN: Cf4  
US-08-945-983-2

Query Match 1.0%; Score 8; DB 4; Length 806;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSSNKI 730  
Db 664 LDLSSNKI 671

RESULT 19  
US-08-666-271-2  
; Sequence 2, Application US/08666271  
; Patent No. 592000  
; GENERAL INFORMATION:  
; APPLICANT: JONES, JONATHAN D  
; APPLICANT: HAMMOND-KOSACK, KIM E  
; APPLICANT: THOMAS, COLWYN M  
; APPLICANT: JONES, DAVID A  
; TITLE OF INVENTION: PLANT PATHOGEN RESISTANCE GENES AND USES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/666,271  
; FILING DATE: 19-SEP-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB94/02812  
; FILING DATE: 23-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9326428.1  
; FILING DATE: 24-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9409363.0  
; FILING DATE: 11-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B.J.  
; REGISTRATION NUMBER: 36,663  
; REFERENCE/DOCKET NUMBER: 620-7  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-816-4091  
; TELEFAX: 703-816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 863 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-666-271-2

Query Match 1.0%; Score 8; DB 2; Length 863;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 723 LDLSSNKI 730  
Db 721 LDLSSNKI 728

RESULT 20  
US-08-475-891A-4

```
; Sequence 4, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1012
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp.
; OTHER INFORMATION: disease resistance polypeptide RRK-B
; OTHER INFORMATION: from rice (Oryza sativa)"
; US-08-475-891A-4

Query Match 1.0%; Score 8; DB 2; Length 1012;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682
Db 156 LDLSHNQL 163

RESULT 21
US-08-475-891A-2
; Sequence 2, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Yuang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
```

```
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1023
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp.
; OTHER INFORMATION: disease resistance polypeptide RRK-F
; OTHER INFORMATION: from rice (Oryza sativa)"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1010
; OTHER INFORMATION: /note= "Xaa = Leu when position
; OTHER INFORMATION: 5471 of RRK-F = G or Phe when position
; OTHER INFORMATION: 5471 of RRK-F = C"
; US-08-475-891A-2

Query Match 1.0%; Score 8; DB 2; Length 1023;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682
Db 166 LDLSHNQL 173

RESULT 22
US-08-567-375-2
; Sequence 2, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/567,375  
FILING DATE: 04-DEC-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,645  
FILING DATE: 29-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/475,891  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,375  
FILING DATE: 17-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-058930  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1023 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1023  
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease  
OTHER INFORMATION: resistance polypeptide RRK-F from rice  
OTHER INFORMATION: RRK-F = G or Phe when position 5471 of  
OTHER INFORMATION: RRK-F = C"  
US-08-567-375-2

Query Match 1.0%; Score 8; DB 2; Length 1023;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682  
Db 166 LDLSHNQL 173

RESULT 23  
US-08-587-680A-2  
Sequence 2, Application US/08587680A  
Patent No. 5977434  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Conferring  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/587,680A  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,375  
FILING DATE: 17-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/475,891  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,645  
FILING DATE: 29-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/567,375  
FILING DATE: 04-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-0589400S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1023 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1023  
OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease  
OTHER INFORMATION: resistance polypeptide RRK-F from rice  
OTHER INFORMATION: RRK-F = G or Phe when position 5471 of  
OTHER INFORMATION: RRK-F = C"  
US-08-587-680A-2

Query Match 1.0%; Score 8; DB 2; Length 1023;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 675 LDLSHNQL 682  
Db 166 LDLSHNQL 173

RESULT 24  
US-08-567-375-4  
Sequence 4, Application US/08567375  
Patent No. 5952485  
GENERAL INFORMATION:  
APPLICANT: Ronald, Pamela C.  
APPLICANT: Wang, Guo-Liang  
APPLICANT: Song, Wen-Yuang  
APPLICANT: Szabo, Veronique  
TITLE OF INVENTION: Procedures and Materials for Conferring

;; TITLE OF INVENTION: Disease Resistance in Plants  
;; NUMBER OF SEQUENCES: 16  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend and Crew LLP  
;; STREET: Two Embarcadero Center, Eighth Floor  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94111-3834  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/567,375  
;; FILING DATE: 04-DEC-1995  
;; CLASSIFICATION: 800  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/373,375  
;; FILING DATE: 17-JAN-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/004,645  
;; FILING DATE: 29-SEP-1995  
;; APPLICATION DATA:  
;; NAME: Bastian, Kevin L.  
;; REGISTRATION NUMBER: 34,774  
;; REFERENCE/DOCKET NUMBER: 023070-058930  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 576-0200  
;; TELEFAX: (415) 576-0300  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1025 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-567-375-4

Query Match 1.0%; Score 8; DB 2; Length 1025;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 675 LDLSHNQL 682  
Db 156 LDLSHNQL 163

RESULT 25  
US-08-587-680A-4  
; Sequence 4, Application US/08587680A  
; Patent No. 5977434  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Conferring  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/587,680A  
;; FILING DATE: 17-JAN-1996  
;; CLASSIFICATION: 800  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/373,375  
;; FILING DATE: 17-JAN-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/475,891  
;; FILING DATE: 07-JUN-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/004,645  
;; FILING DATE: 29-SEP-1995  
;; APPLICATION DATA:  
;; NAME: Bastian, Kevin L.  
;; REGISTRATION NUMBER: 34,774  
;; REFERENCE/DOCKET NUMBER: 023070-0589400S  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 576-0200  
;; TELEFAX: (415) 576-0300  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1025 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-587-680A-4

Query Match 1.0%; Score 8; DB 2; Length 1025;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 675 LDLSHNQL 682  
Db 156 LDLSHNQL 163

RESULT 26  
US-08-710-749-13  
; Sequence 13, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.



REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-13

Query Match 0.9%; Score 7; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 620 LKLEEL 626  
DB 32 LKLEEL 38

RESULT 27  
US-08-710-749-14  
; Sequence 14, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-14

Query Match 0.9%; Score 7; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 620 LKLEEL 626

DB 32 LKLEEL 38

RESULT 28  
US-08-710-749-16  
; Sequence 16, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-16

Query Match 0.9%; Score 7; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 620 LKLEEL 626  
DB 32 LKLEEL 38

RESULT 29  
US-08-710-749-8  
; Sequence 8, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York

```
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/710,749
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-710-749-8

; Query Match 0.9%; Score 7; DB 2; Length 102;
; Best Local Similarity 100.0%; Pred. No. 68;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 620 LKLEEL 626
Db 33 LKLEEL 39

RESULT 30
US-09-228-986-99
; Sequence 99, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-99

; Query Match 0.9%; Score 7; DB 4; Length 154;
; Best Local Similarity 100.0%; Pred. No. 97;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 315 QLDLSEF 321
Db 121 QLDLSEF 127

RESULT 31
US-09-228-986-99
; Sequence 90, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-228-986-99

; Query Match 0.9%; Score 7; DB 4; Length 157;
; Best Local Similarity 100.0%; Pred. No. 98;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 542 LDISSNS 548
Db 126 LDISSNS 132

RESULT 32
US-09-117-257-38
; Sequence 38, Application US/09117257
; Patent No. 6214355
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; APPLICANT: Hanson, Mark
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 4210 000500
; CURRENT APPLICATION NUMBER: US/09/117,257
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: PCT/US96/17081
; EARLIER FILING DATE: 1996-10-22
; EARLIER APPLICATION NUMBER: 08/589,711
; EARLIER FILING DATE: 1996-01-22
; EARLIER APPLICATION NUMBER: 08/427,023
; EARLIER FILING DATE: 1995-04-24
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 38
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-117-257-38

; Query Match 0.9%; Score 7; DB 4; Length 160;
; Best Local Similarity 100.0%; Pred. No. 1e+02;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 FKLLKL 623
Db 13 FKLLKL 19

RESULT 33
US-09-489-352-38
; Sequence 38, Application US/09489352
; Patent No. 6312907
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; APPLICANT: Hanson, Mark
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 4210 000500
; CURRENT APPLICATION NUMBER: US/09/489,352
; CURRENT FILING DATE: 2000-01-21
; EARLIER APPLICATION NUMBER: PCT/US96/17081
```

; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 38  
; LENGTH: 160  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-489-352-38

Query Match 0.9%; Score 7; DB 4; Length 160;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 FKLLKL 623  
Db 13 FKLLKL 19

RESULT 34  
US-08-923-738-2  
; Sequence 2, Application US/08923738A  
; Patent No. 5948642  
; GENERAL INFORMATION:  
; APPLICANT: Burnham, Martin K. R.  
; APPLICANT: Lonetto, Michael A.  
; APPLICANT: Warren, Patrick V.  
; TITLE OF INVENTION: No. 5948642el DNA Strand Resolution  
; FILE REFERENCE: P50549-06  
; CURRENT APPLICATION NUMBER: US/08/923,738A  
; CURRENT FILING DATE: 1997-09-02  
; EARLIER APPLICATION NUMBER: 60/027,032  
; EARLIER FILING DATE: 1996-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 172  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-08-923-738-2

Query Match 0.9%; Score 7; DB 2; Length 172;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 LEVLDLG 376  
Db 89 LEVLDLG 95

RESULT 35  
US-08-923-738-4  
; Sequence 4, Application US/08923738A  
; Patent No. 5948642  
; GENERAL INFORMATION:  
; APPLICANT: Burnham, Martin K. R.  
; APPLICANT: Lonetto, Michael A.  
; APPLICANT: Warren, Patrick V.  
; TITLE OF INVENTION: No. 5948642el DNA Strand Resolution  
; FILE REFERENCE: P50549-06  
; CURRENT APPLICATION NUMBER: US/08/923,738A  
; CURRENT FILING DATE: 1997-09-02  
; EARLIER APPLICATION NUMBER: 60/027,032  
; EARLIER FILING DATE: 1996-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 172

; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-08-923-738-4

Query Match 0.9%; Score 7; DB 2; Length 172;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 LEVLDLG 376  
Db 89 LEVLDLG 95

RESULT 36  
US-08-936-165A-385  
; Sequence 385, Application US/08936165A  
; Patent No. 6348582  
; GENERAL INFORMATION:  
; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Lonetto, Michael  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Pratt, Julie  
; APPLICANT: Reichard, Richard  
; APPLICANT: Rosenberg, Martin  
; APPLICANT: Ward, Judith  
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,  
; TITLE OF INVENTION: Polypeptides and Their Uses  
; NUMBER OF SEQUENCES: 534  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/936,165A  
; FILING DATE: 24-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/027,032  
; FILING DATE: 24-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gimmli, Edward R  
; REGISTRATION NUMBER: 38,891  
; REFERENCE/DOCKET NUMBER: P50549  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-4478  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 385:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 172 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
US-08-936-165A-385

Query Match 0.9%; Score 7; DB 4; Length 172;  
Best Local Similarity 100.0%; Pred. No. 1.le+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 370 LEVLDLG 376  
|  
Db 92 LEVLDLG 98

RESULT 37  
US-09-475-316A-23  
; Sequence 23, Application US/09475316A  
; Patent No. 6210942  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, No. 6210942man G.  
; APPLICANT: Davis, Laurence B.  
; APPLICANT: Dinkova-Kostova, Albena T.  
; APPLICANT: Fujita, Masayuki  
; APPLICANT: Gang, David R.  
; APPLICANT: Sarkanen, Simo  
; APPLICANT: Ford, Joshua D.  
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICIRSINOL REDUCTASES,  
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE  
; FILE REFERENCE: WSUR-1-13793  
; CURRENT FILING DATE: 1999-12-30  
; CURRENT FILING DATE: 1999-12-30  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: PCT/US97/20391  
; PRIOR FILING DATE: 1997-11-07  
; PRIOR APPLICATION NUMBER: 60/054,380  
; PRIOR FILING DATE: 1997-07-31  
; PRIOR APPLICATION NUMBER: 60/030,522  
; PRIOR FILING DATE: 1996-11-08  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Thuja plicata  
US-09-475-316A-23

Query Match 0.9%; Score 7; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 AKFLHFL 310  
|  
Db 6 AKFLHFL 12

RESULT 38  
US-09-117-257-17  
; Sequence 17, Application US/09117257  
; Patent No. 6214355  
; GENERAL INFORMATION:  
; APPLICANT: Hook, Magnus  
; APPLICANT: Guo, Betty  
; APPLICANT: Hanson, Mark  
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: 4210,000500  
; CURRENT FILING DATE: 1998-07-22  
; CURRENT FILING DATE: 1998-07-22  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-117-257-17

Query Match 0.9%; Score 7; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 FKLLKL 623  
|  
Db 13 FKLLKL 19

RESULT 39  
US-08-945-476-17  
; Sequence 17, Application US/08945476  
; Patent No. 6248517  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS OF USE  
; NUMBER OF SEQUENCES: 27  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/945,476  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/589,711  
; FILING DATE: 22-JAN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/427,023  
; FILING DATE: 24-APR-1995  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 194 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-945-476-17

Query Match 0.9%; Score 7; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 FKLLKL 623  
|  
Db 13 FKLLKL 19

RESULT 40  
US-09-489-352-17  
; Sequence 17, Application US/09489352  
; Patent No. 6312907  
; GENERAL INFORMATION:  
; APPLICANT: Hook, Magnus  
; APPLICANT: Guo, Betty  
; APPLICANT: Hanson, Mark  
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: 4210,000500  
; CURRENT FILING DATE: 2000-01-21  
; CURRENT FILING DATE: 2000-01-21  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17

; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-489-352-17

Query Match 0.9%; Score 7; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 EKNLLKL 623  
Db 13 FRNLLKL 19

## RESULT 41

US-09-117-257-34  
; Sequence 34, Application US/09117257  
; Patent No. 6214355  
; GENERAL INFORMATION:  
; APPLICANT: Hook, Magnus  
; APPLICANT: Guo, Betty  
; APPLICANT: Hanson, Mark  
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: 4210.000500  
; CURRENT APPLICATION NUMBER: US/09/117,257  
; CURRENT FILING DATE: 1998-07-22  
; EARLIER APPLICATION NUMBER: PCT/US96/17081  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-117-257-34

Query Match 0.9%; Score 7; DB 4; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 EKNLLKL 623  
Db 13 FRNLLKL 19

## RESULT 42

US-09-117-257-46  
; Sequence 46, Application US/09117257  
; Patent No. 6214355  
; GENERAL INFORMATION:  
; APPLICANT: Hook, Magnus  
; APPLICANT: Guo, Betty  
; APPLICANT: Hanson, Mark  
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: 4210.000500  
; CURRENT APPLICATION NUMBER: US/09/117,257  
; CURRENT FILING DATE: 1998-07-22  
; EARLIER APPLICATION NUMBER: PCT/US96/17081  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 195  
; TYPE: PRT

; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-117-257-46

Query Match 0.9%; Score 7; DB 4; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 EKNLLKL 623  
Db 13 FRNLLKL 19

## RESULT 43

US-09-489-352-34  
; Sequence 34, Application US/09489352  
; Patent No. 6312907  
; GENERAL INFORMATION:  
; APPLICANT: Hook, Magnus  
; APPLICANT: Guo, Betty  
; APPLICANT: Hanson, Mark  
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: 4210.000500  
; CURRENT APPLICATION NUMBER: US/09/489,352  
; CURRENT FILING DATE: 2000-01-21  
; EARLIER APPLICATION NUMBER: PCT/US96/17081  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Borrelia burgdorferi  
US-09-489-352-34

Query Match 0.9%; Score 7; DB 4; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 EKNLLKL 623  
Db 13 FRNLLKL 19

## RESULT 44

US-09-489-352-46  
; Sequence 46, Application US/09489352  
; Patent No. 6312907  
; GENERAL INFORMATION:  
; APPLICANT: Hook, Magnus  
; APPLICANT: Guo, Betty  
; APPLICANT: Hanson, Mark  
; TITLE OF INVENTION: DbpA AND DbpB COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: 4210.000500  
; CURRENT APPLICATION NUMBER: US/09/489,352  
; CURRENT FILING DATE: 2000-01-21  
; EARLIER APPLICATION NUMBER: PCT/US96/17081  
; EARLIER FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/589,711  
; EARLIER FILING DATE: 1996-01-22  
; EARLIER APPLICATION NUMBER: 08/427,023  
; EARLIER FILING DATE: 1995-04-24  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 195  
; TYPE: PRT

; ORGANISM: Borrelia burgdorferi  
US-09-489-352-46

Query Match 0.9%; Score 7; DB 4; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 FKNLLKL 623  
Db 13 FKNLLKL 19  
|||||||

## RESULT 45

US-09-228-986-93  
; Sequence 93, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 93  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-93

Query Match 0.9%; Score 7; DB 4; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 YLDLSSN 728  
Db 145 YLDLSSN 151  
|||||||

## RESULT 46

US-08-211-682-25  
; Sequence 25, Application US/08211682  
; Patent No. 5670333  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: EXPRESSION OF POLYPEPTIDES IN E.COLI UNDER  
; TITLE OF INVENTION: CONTROL OF THE E.COLI MDH-GENE PROMOTER  
; NUMBER OF SEQUENCES: 25  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/211,682  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 327 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-211-682-25

Query Match 0.9%; Score 7; DB 1; Length 327;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 QLLEIPQ 115  
Db 39 QLLEIPQ 45  
|||||||

## RESULT 47

US-09-091-405-2  
; Sequence 2, Application US/09091405A  
; Patent No. 6207144  
; GENERAL INFORMATION:  
; APPLICANT: KORTH, Reinhard  
; APPLICANT: BAIER, Michael  
; APPLICANT: BANNERT, No. 6207144bert  
; APPLICANT: METZNER, Karin  
; APPLICANT: WERNER, Albrecht  
; APPLICANT: LANG, Kurt  
; TITLE OF INVENTION: POLYPEPTIDES WITH INTERLEUKIN 16 ACTIVITY, PROCESS FOR  
; TITLE OF INVENTION: THE PREPARATION AND USE THEREOF  
; FILE REFERENCE: 8341-8041  
; CURRENT APPLICATION NUMBER: US/09/091,405A  
; CURRENT FILING DATE: 1998-08-27  
; EARLIER APPLICATION NUMBER: PCT/EP96/05662  
; EARLIER FILING DATE: 1996-12-17  
; EARLIER APPLICATION NUMBER: DE/196 13 886.8  
; EARLIER FILING DATE: 1996-04-06  
; EARLIER APPLICATION NUMBER: DE/196 13 866.3  
; EARLIER FILING DATE: 1996-04-06  
; EARLIER APPLICATION NUMBER: DE/196 03 492.2  
; EARLIER FILING DATE: 1996-01-31  
; EARLIER APPLICATION NUMBER: DE/195 48 295.6  
; EARLIER FILING DATE: 1995-12-22  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: IL-16  
US-09-091-405-2

Query Match 0.9%; Score 7; DB 4; Length 374;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 GQSVISL 799  
Db 120 GQSVISL 126  
|||||||

## RESULT 48

US-08-525-697-2  
; Sequence 2, Application US/08525697  
; Patent No. 5795764  
; GENERAL INFORMATION:  
; APPLICANT: Christgau, Stephan  
; APPLICANT: Andersen, Lene N  
; APPLICANT: Kauppinen, Sakari  
; APPLICANT: Heldt-Hansen, Hans P  
; APPLICANT: Dalboege, Henrik  
; TITLE OF INVENTION: AN ENZYME EXHIBITING MANNANASE ACTIVITY  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5795764o No. 5795764disk of No. 5795764th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; Zip: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,697  
FILING DATE: 21-SEP-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Harrington, James J.  
REGISTRATION NUMBER: 38,711  
REFERENCE/DOCKET NUMBER: 4004.204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DK 0486/93  
FILING DATE: 30-APR-1993  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-525-697-2

Query Match 0.9%; Score 7; DB 1; Length 377;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 LNFTKNL 566  
|||||||  
DB 254 LNFTKNL 260

RESULT 49  
US-08-775-428-2  
Sequence 2, Application US/08775428  
Patent No. 5976834  
GENERAL INFORMATION:  
APPLICANT: Sathe, Ganesh  
APPLICANT: Fuetterer, Wendy  
APPLICANT: Bergsma, Derk  
APPLICANT: Ellis, Catherine  
TITLE OF INVENTION: CDNA CLONE HNFJD15 THAT ENCODES  
TITLE OF INVENTION: A NOVEL HUMAN 7-TRANSMEMBRANE RECEPTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/775,428  
FILING DATE: 09-JAN-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Han, William T  
REGISTRATION NUMBER: 34,344  
REFERENCE/DOCKET NUMBER: ATG50042  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5219  
TELEFAX: 610-270-4060

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 396 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-775-428-2

Query Match 0.9%; Score 7; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 GDSSEVG 414  
|||||||  
DB 269 GDSSEVG 275

RESULT 50  
US-09-461-474-10  
Sequence 10, Application US/09461474  
Patent No. 6278042  
GENERAL INFORMATION:  
APPLICANT: Allen, Steve  
APPLICANT: Rafalski, Antoni  
APPLICANT: Sakai, Hajime  
TITLE OF INVENTION: Plant Metal Transporters  
FILE REFERENCE: BBI303 US NA  
CURRENT APPLICATION NUMBER: US/09/461,474  
CURRENT FILING DATE: 1999-12-14  
EARLIER APPLICATION NUMBER: 60/1112,562  
EARLIER FILING DATE: 1998-12-16  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 10  
LENGTH: 474  
TYPE: PRT  
ORGANISM: Oryza sativa  
US-09-461-474-10

Query Match 0.9%; Score 7; DB 4; Length 474;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 578 DISSSTS 584  
|||||||  
DB 73 DISSSTS 79

Search completed: July 17, 2002, 09:44:49  
Job time: 41 sec









/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU55027 - 7.00 105.06 1.8e+03 59  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB28287 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAB33467 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB18927 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM54244 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM66637 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM14510 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM26927 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM23236 + 7.00 104.95 1.8e+03 60  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU52342 + 7.00 104.85 1.8e+03 61  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM86744 + 7.00 104.85 1.8e+03 61  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU17428 + 7.00 104.85 1.8e+03 61  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU64288 + 7.00 104.74 1.8e+03 62  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM79753 + 7.00 104.74 1.8e+03 62  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM77747 + 7.00 104.74 1.8e+03 62  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AA59759 + 7.00 104.64 1.8e+03 63  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAG551710 + 7.00 104.64 1.8e+03 63  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM37666 + 7.00 104.64 1.8e+03 63  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU64001 + 7.00 104.54 1.8e+03 64  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AAY41316 + 7.00 104.44 1.8e+03 65  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAB16408 + 7.00 104.44 1.8e+03 65  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG55104 + 7.00 104.34 1.8e+03 66  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAG59751 + 7.00 104.34 1.8e+03 66  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG55283 + 7.00 104.25 1.8e+03 67  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU59704 + 7.00 104.06 1.7e+03 69  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM71887 + 7.00 103.96 1.7e+03 70  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG55123 + 7.00 103.96 1.7e+03 70  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG60656 + 7.00 103.96 1.7e+03 70  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAG551710 + 7.00 103.78 1.7e+03 72  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU62461 + 7.00 103.70 1.7e+03 73  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAO08330 + 7.00 103.61 1.7e+03 74  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAO00358 + 7.00 103.52 1.7e+03 75  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB44152 + 7.00 103.44 1.7e+03 76  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB27034 + 7.00 103.44 1.7e+03 76  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM65179 + 7.00 103.44 1.7e+03 76  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM7786 + 7.00 103.44 1.7e+03 76  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM21781 + 7.00 103.44 1.7e+03 76  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM38106 + 7.00 103.44 1.7e+03 76  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AAM7876 + 7.00 103.27 1.7e+03 78  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABG02029 + 7.00 103.27 1.7e+03 78  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABG18777 + 7.00 103.27 1.7e+03 78  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG15015 + 7.00 103.11 1.7e+03 80  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM92606 + 7.00 103.11 1.7e+03 80  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAO07445 + 7.00 103.11 1.7e+03 80  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAB38613 + 7.00 102.95 1.7e+03 82  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU45065 + 7.00 102.95 1.7e+03 82  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU63632 + 7.00 102.80 1.7e+03 84  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU59472 + 7.00 102.65 1.7e+03 86  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU0108 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB41823 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB25543 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM62695 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM75513 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM20619 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAM35619 + 7.00 102.57 1.7e+03 87  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU28364 + 7.00 102.43 1.7e+03 89  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.AAG80079 + 7.00 102.35 1.7e+03 90  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG54723 + 7.00 102.35 1.7e+03 90  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG42147 + 7.00 102.28 1.7e+03 91  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG641708 + 7.00 102.28 1.7e+03 91  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU48773 + 7.00 102.28 1.7e+03 91  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAU62952 + 7.00 102.28 1.7e+03 91  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB17526 + 7.00 102.28 1.7e+03 91  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAO06581 + 7.00 102.28 1.7e+03 91  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG09848 + 7.00 102.21 1.7e+03 92  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG50982 + 7.00 102.21 1.7e+03 92  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAG74526 + 7.00 102.21 1.7e+03 92  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAG75361 + 7.00 102.21 1.7e+03 92  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAO04793 + 7.00 102.14 1.7e+03 93  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AAM78227 + 7.00 102.08 1.7e+03 94  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAG57612 + 7.00 102.08 1.7e+03 94  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAG02085 + 7.00 102.08 1.7e+03 94  
/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABG04984 + 7.00 102.08 1.7e+03 94















CC disorders. They may also be used to raise antibodies. AA233891 to  
CC AA234338, and AA41685 to AA41774 represent polynucleotide and  
CC polypeptide sequence given in the exemplification of the present  
CC invention.

XX Sequence 1049 AA;

alignment\_scores:

Quality: 1049.00 Length: 1049  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AA41767 ..

Align seg 1/1 to: AA41767 from: 1 to: 1049

```
85 ATGGTGTTCCTCAATGGGACATGAGACAAATCTTATCCTTTTAA 134
|||||
1 MetValPheProMetTrpThrLeuLysArgGlnLeuLeuPheAs 17
|||||
135 CATAAATCCTAATTCCTCAACTCTGGGCTAGATGGTTCCCTAAACCTC 184
|||||
17 nileLeuLeuSerLysLeuLeuGlyAlaArgTrpPheProLysThrL 34
|||||
185 TGGCCTGTGATGTCACTCTGGATGTTCCTCAAGAACCATGTGATCGTGGAC 234
|||||
34 euProCysAspValThrLeuAspValProLysAsnHisValIleValAsp 50
|||||
235 TGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCTCCACGACAC 284
|||||
51 CysThrAspLysHisLeuThrGluIleProGlyGlyLeuProThrAsnTh 67
|||||
285 CAGAACTCACCCTCACCATTAAACACATACCAGACATCTCCCGACCGT 334
|||||
67 rThrAsnLeuThrLeuThrIleAsnHisIleProAspIleSerProAla 84
|||||
335 CCTTTCACAGACTGGACCTCTGGTAGAGATCGATTTCAGATCAACTGT 384
|||||
84 erPheHisArgLeuAspHisLeuValGluIleAspPheArgCysAsnCys 100
|||||
385 GTACCTATTCCACTGGGGTCAAAACAACAACATGTGCATCAAGAGGCTGCA 434
|||||
101 ValProIleProLeuGlySerLysAsnMetCysIleLysArgLeuGlu 117
|||||
435 GATTAAACCCAGAGCTTTAGTGGACTCACTTATTTAAATCCCTTACC 484
|||||
117 nileLysProArgSerPheSerGlyLeuThrTyrLeuLysSerLeuTyrL 134
|||||
485 TGGATGGAACCCAGCTACTAGAGATACCGAGGCGCTCCCGCCTAGCTTA 534
|||||
134 euAspGlyAsnGlnLeuLeuGluIleProGlnGlyLeuProProSerLeu 150
|||||
535 CAGCTTCTCAGCCTTGAGGCCAACAAACATCTTTCCATCAGAAAGAGAA 584
|||||
151 GlnLeuLeuSerLeuGluAlaAsnAsnIlePheSerIleArgLysGluAs 167
|||||
585 TCTAACAGAACTGCCCAACATAGAAATACTCTACCTGGGGCCAAACTGTT 634
|||||
167 nLeuThrGluLeuAlaAsnIleGluIleLeuTyrLeuGlyGlnAsnCys 184
|||||
635 ATTATCGAAATCCTGTGTATGTTCATATTCAATAGAGAAGATGCTTTC 684
|||||
184 yrTyrArgAsnProCysTyrValSerTyrSerIleGluLysAspAlaPhe 200
|||||
685 CTAACCTTGACAAAGTTAAAGTCTCTCCCTGAAAGATAAACATGTTCAC 734
|||||
201 LeuAsnLeuThrLysLeuLysValLeuSerLeuLysAspAsnValTh 217
|||||
735 AGCGGTCCTACTGTTTGGCATCTACTTTAACAGAACTATATCTCTACA 784
|||||
217 rAlaValProThrValLeuProSerThrThrLeuThrGluLeuTyrLeuTyr 234
|||||
```

```
785 ACAACATGATTGCAAAATCCAAAGAGATGATTTTAAATACCTCAACCAA 834
|||||
234 snAsnMetIleAlaLysIleGlnGluAspPheAsnAsnLeuAsnGln 250
|||||
835 TTACAAATCTTTGACCTAAGTGGAAATGCTCCTCGTTGTATTAATGCCCC 884
|||||
251 LeuGlnIleLeuAspLeuSerGlyAsnCysProArgCysTyrAsnAlaPr 267
|||||
885 ATTTCTCTCTGGCGCGTGTAAATAATCTCCCTACAGATCCCTGTAA 934
|||||
267 opheProCysAlaProCysLysAsnAsnSerProLeuGlnIleProValA 284
|||||
935 ATGCTTTTGTGCTGACAGATTAATAAGTTTACGCTCTACACAGTAAC 984
|||||
284 snAlaPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300
|||||
985 TCTCTTCAGCATGTGCCCCCAAGATGTTTAAGAACATCAACAACTCCA 1034
|||||
301 SerLeuGlnHisValProArgTrpPheLysAsnIleAsnLysLeuGlu 317
|||||
1035 GGAACCTGGATCTGTCCCAAACTTCTTGCCCAAGAAATTTGGGATGCTA 1084
|||||
317 nGluLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaL 334
|||||
1085 AATTCTCTGCATTTTCTCCAGCCTCATCAATTTGGATCTGTCTTCAAT 1134
|||||
334 yspPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350
|||||
1135 TTTGAACTTCAGGCTCATCTGTCATCTATGAATCTATCACAAGCATTTTC 1184
|||||
351 PheGluLeuGlnValTyrArgAlaSerMetAsnLeuSerGlnAlaPheSe 367
|||||
1185 TTCACCTGAAAGCCTGAAATTTCTGCGGATCAGAGGATATGCTTTAAAG 1234
|||||
367 rSerLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysG 384
|||||
1235 AGTTGAAAAGCTTTAACTCTCGCCATTACATAATCTTCAAAATCTTGAA 1284
|||||
384 luLeuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400
|||||
1285 GTTCTTGATCTTGGCAGCTAACTTTATAAATTTGCTAACCTCAGCATGTT 1334
|||||
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417
|||||
1335 TAAACAATTTAAAGACTGAAACTCATAGATCTTCAGTGAATAAATAT 1384
|||||
417 euLysGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIle 434
|||||
1385 CACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAATGCCAGAACT 1434
|||||
434 erProSerGlyAspSerSerGluValGlyPheCysSerAsnAlaArgThr 450
|||||
1435 TCTGTGAAAGTTATGAACCCAGGCTCGGAACAATTACATATTTCAG 1484
|||||
451 SerValGluSerTyrGluProGlnValLeuGluGlnLeuHisTyrPheAr 467
|||||
1485 ATATGATAGATATGCAAGGAGTTGCAGATTCAAACAAAGAGGCTCTTT 1534
|||||
467 gTyrAspLysTyrAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484
|||||
1535 TCATGTCTGTTAATGAAAGCTGTACAAGTATGGCAGACCTTTGGATCTA 1584
|||||
484 heMetSerValAsnGluSerCysTyrLysTyrGlyGlnThrLeuAspLeu 500
|||||
1585 AGTAAATAAGTATATTTTGTCAAGTCTCTGATTTTCAGCATCTTTC 1634
|||||
501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517
|||||
1635 TTTTCCCTCAAAATGCTGAATCTGTGAGAAATCTCATTAGCCAAACTCTTA 1684
|||||
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534
|||||
```

1685 ATGCAGTGAATCCAACTTTAGCAGAGCTGAGATATTGGACTTCTCC 1734  
|||||  
534 snglyserglupheglnproleualagluLeuArgTyrLeuAspPheSer 550  
1735 AACACCGCGTGTGATTACTCCATTCAACAGCATTTGAAGAGCTTCACAA 1784  
|||||  
551 AsnAsnArgLeuAspLeuLeuHisSerThrAlaPheGluGluLeuHisLy 567  
1785 ACTGGAAGTCTGGATATAGCGAGTATAGCCATTATTTCATCAGAAG 1834  
|||||  
567 sLeuGluValLeuAspPheSerAsnSerHisTyrPheGlnSerGluG 584  
1835 GAATTACTCATATGCTAACTTTACCAAGAACTTAAAGGTTCTGCAGAAA 1884  
|||||  
584 lyleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600  
1885 CTGATGATGACGACATGACATCTCTTCTCCACGACGAGGACCATGGA 1934  
|||||  
601 LeuMetMetAsnAspAsnAspIleSerSerThrSerArgThrMetG1 617  
1935 GAGTGACTCTTAGAACTCTGGAATTCAGAGGAATCACTTAGATGTTT 1984  
|||||  
617 userGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspVal 634  
1985 TATGGAGAGAGGTGATAACAGATACTTACAATTATTCAAGAATCTGCTA 2034  
|||||  
634 euTrpArgGluGlyAspAsnArgTyrLeuGlnLeuPheLysAsnLeuLeu 650  
2035 AAATTAGAGGAATTAGACATCTCTAAAATTCCTTAAGTTTCTTGCCCTC 2084  
|||||  
651 LysLeuGluGluLeuAspIleSerLysAsnSerLeuSerPheLeuProSe 667  
2085 TGGAGTTTGTGAGTATGCTCCAAATCTAAAGAATCTCTCTTTGGCCA 2134  
|||||  
667 rGlyValPheAspGlyMetProAsnLeuLysAsnLeuSerLeuAlaL 684  
2135 AAAATGGGCTCAATCTTTTCAGTTGGAAGAACTCCAGTGTCTAAAGAAC 2184  
|||||  
684 ysAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
2185 CTGGAATCTTGGACCTCACCACCAACTGACCACTGTCCTCTGAGAG 2234  
|||||  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrValProGluAr 717  
2235 ATTATCCAATCTTCCAGAAGCTCAAGAATCTGATTCTTTAAGAATAATC 2284  
|||||  
717 gLeuSerAsnCysSerArgSerLeuLysAsnLeuLeuLeuLysAsnSng 734  
2285 AAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCTTCCAGTTGCCGA 2334  
|||||  
734 lnIleArgSerLeuThrLysTy:PheLeuGlnAspAlaPheGlnLeuArg 750  
2335 TATCTGATCTCAGCTCAATATAATCCAGATGATCCAAAAGACCAGCTT 2384  
|||||  
751 TyrLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
2385 CCCAGAAAATGCTCTCAACAATCTGAAGATGTTGCTTTTCATCATATAATC 2434  
|||||  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuLeuHisHisAsnA 784  
2435 GGTTCCTGTCACCTGTGATGCTGTGGTTGTGCTGGTGGGTAAACCAT 2484  
|||||  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpTrpValAsnHis 800  
2485 ACGGAGGTGACTATTCTTACCTGGCCACAGATGCTACTTGTGTGGGGCC 2534  
|||||  
801 ThrGluValThrIleProTyrLeuAlaThrAspValThrCysValGlyPr 817  
2535 AGGAGCACACAAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
|||||  
817 OGlyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyrThrCysG 834  
2585 AGTTAGATCTGACTAACTGATCTGTGTTCTCTCACTTTTCCATATCTGTATCT 2634

|||||  
834 luLeuAspLeuThrAsnLeuLeuPheSerLeuSerIleSerValSer 850  
2635 CTCTTCTCATGTGATGATGACAGCAAGTCACTCTATTCTTCTGGGATGT 2684  
|||||  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyrPheTrpAspVa 867  
2685 GTGGTATATTTACCATTCTCTAAGCCCAAGATAAAAGGGGTATCAGCGTC 2734  
|||||  
867 lTrpTyrIleTyrHisPheCysLysAlaLysIleLysGlyTyrGlnArgL 884  
2735 TAATATCACCAGACTGTTGCTGATGCTTTTATTGTGTATGACACTAAA 2784  
|||||  
884 euIleSerProAspCysCysTyrAspAlaPheIleValTyrAspThrLys 900  
2785 GACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGGTGGCCCAACTGGA 2834  
|||||  
901 AspProAlaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuG1 917  
2835 AGACCCAAAGAGAGAAACATTTTAATTTATGCTCTCGAGGAAGGACTGGT 2884  
|||||  
917 uAspProArgGluLysHisPheAsnLeuCysLeuGluGluArgAspTrpL 934  
2885 TACCAGGCGCAGCCAGTCTTGAAAACTTTCCAGAGCATACAGCTTAGC 2934  
|||||  
934 euProGlyGlnProValLeuGluAsnLeuSerGlnSerIleGlnLeuSer 950  
2935 AAAAAGACAGTGTGTGATGACAGACAGATGTCGAAGACTGAAAATTT 2984  
|||||  
951 LysLysThrValPheValMetThrAspLysTyrAlaLysThrGluAsnPh 967  
2985 TAAGTAGCATTCTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTG 3034  
|||||  
967 eLysIleAlaPheTyrLeuSerHisGlnArgLeuMetAspGluLysVala 984  
3035 ATGTGATTATCTTGTATATTTCTTGAGAAGCCCTTTGAGAAGTCCAAGTTC 3084  
|||||  
984 spValIleIleLeuIlePheLeuGluLysProPheGlnLysSerLysPhe 1000  
3085 CTCCAGCTCCGGAAAAAGGCTCTGTGGAGTTCCTCTTGTAGTGGCCAAC 3134  
|||||  
1001 LeuGlnLeuArgLysArgLeuCysGlySerSerValLeuGluTrpProTh 1017  
3135 AAACCCGCAAGCTCACCATCTTCTGGCAGTGTCTAAAGAACCCCTCG 3184  
|||||  
1017 rAsnProGlnAlaHisProTyrPheTrpGlnCysLeuLysAsnAlaLeuA 1034  
3185 CCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAAACCGTC 3231  
|||||  
1034 laThrAspAsnHisValAlaTyrSerGlnValPheLysGluThrVal 1049  
seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA1999

seq\_documentation\_block:  
ID AAY05866 standard; protein; 1049 AA.

XX AAY05866;

XX 02-AUG-1999 (first entry)

XX Human Toll protein PRO285.

XX PRO285; Toll; homologue; human; adaptive immunity; septic shock;

KW inflammation; diabetes; amyotrophic lateral sclerosis; cancer;

KW ulcer; rheumatoid arthritis; pathogen pattern recognition receptor;

XX signal transduction.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..29

FT Protein 30..1049

FT Domain /note= "mature protein"  
FT 837..860  
FT /note= "transmembrane domain"  
FT 132..153  
FT Region /note= "leucine zipper"  
FT 704..725  
FT Region /note= "leucine zipper"  
FT 66  
FT Modified-site /note= "N-glycosylated"  
FT 69  
FT Modified-site /note= "N-glycosylated"  
FT 167  
FT Modified-site /note= "N-glycosylated"  
FT 202  
FT Modified-site /note= "N-glycosylated"  
FT 215  
FT Modified-site /note= "N-glycosylated"  
FT 361  
FT Modified-site /note= "N-glycosylated"  
FT 413  
FT Modified-site /note= "N-glycosylated"  
FT 488  
FT Modified-site /note= "N-glycosylated"  
FT 523  
FT Modified-site /note= "N-glycosylated"  
FT 534  
FT Modified-site /note= "N-glycosylated"  
FT 590  
FT Modified-site /note= "N-glycosylated"  
FT 679  
FT Modified-site /note= "N-glycosylated"  
FT 720  
FT Modified-site /note= "N-glycosylated"  
FT 799  
FT Modified-site /note= "N-glycosylated"  
FT 942  
FT Modified-site /note= "N-glycosylated"

W09920756-A2.

29-APR-1999.

07-OCT-1998; 98WO-US21141.

26-JUN-1998; 98US-0105413.  
17-OCT-1997; 97US-0062250.  
13-NOV-1997; 97US-0065311.  
28-APR-1998; 98US-0083322.  
26-JUN-1998; 98US-0090863.

(GETH ) GENENTECH INC.

Goddard A, Godowski PJ, Gurney AL, Mark MR., Yang R;

WPI; 1999-302739/25.

DR N-PSDB; AAX58295.

New human Toll-like receptors that recognize microbial structures

Claim 1; Fig 1; 79pp; English.

XX This is the amino acid sequence of PRO285, a novel human homologue  
CC of Drosophila Toll protein, that acts as a pathogen pattern  
CC recognition receptor, sensing the presence of conserved molecular  
CC structures present on microorganisms. The sequence was deduced  
CC from isolated cDNA clone DNA40021-1154 (ATCC 209389) (see AAX58295).  
CC The invention provides 3 novel cDNA clones that encode novel human  
CC Toll polypeptides PRO285, PRO286 (see AAY05867) and PRO358 (see  
CC AAY05868). It also provides specific antibodies and chimeric  
CC molecules comprising a PRO285, PRO286 or PRO358 polypeptide, or a  
CC transmembrane domain-deleted or inactivated variant, fused to a  
CC heterologous amino acid sequence, such as an epitope tag or  
CC immunoglobulin Fc region. Being homologues of Drosophila Toll

CC protein, the 3 human proteins are likely to be involved in adaptive  
CC immunity, particularly inflammation, septic shock, and response to  
CC pathogens in diseases aggravated by the immune response, such as  
CC diabetes, amyotrophic lateral sclerosis, cancer, rheumatoid  
CC arthritis and ulcers. The PRO polypeptides are used to identify  
CC other proteins involved in Toll-mediated signal transduction (e.g.  
CC natural ligands), to screen for receptor and ligand mimics, and to  
CC generate antibodies. Antibodies specific for the PRO polypeptides  
CC (or for the known receptor TLR-2) are used to treat septic shock  
CC (claimed).

XX Sequence 1049 AA;

alignment\_scores:

Quality: 1049.00 Length: 1049  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAY05866 ..

Align seg 1/1 to: AAY05866 from: 1 to: 1049

85 ATGGTGTTCCTCAATGTGGACACTGAAGACAAATTCCTTTTAA 134  
|||||  
1 MetValPheProMetTrpThrLeuLysArgGlnIleLeuIleLeuPheAs 17  
  
135 CATAATCTTAATTCCTGGGCTAGATGGTTCTCTAAACATC 184  
|||||  
17 nlelleleuIleSerLysLeuGlyAlaArgTrpPheProLysThrL 34  
  
185 TGCCTGTGATGTCACCTCTGGATGTTCCAAAGAACCATCGTGGAC 234  
|||||  
34 euProCysaspValThrLeuaspValProLysasnHisValIleValasp 50  
  
235 TGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCACAGAACAC 284  
|||||  
51 CysThrAspLysHisLeuThrGluIleProGlyGlyIleProThrAsnTh 67  
  
285 CAGGACCTCACCCCTCACCATTAAACACATACCAGACATCTCCACGCGT 334  
|||||  
67 rThrAsnLeuThrLeuThrIleAsnHisIleProaspIleSerProAlas 84  
  
335 CCTTTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGATGCAACTGT 384  
|||||  
84 erPheHisargLeuaspHisLeuValGluIleaspPheargCysasnCys 100  
  
385 GTACCTATTTCCTACTGGGGTCAAAAACAAACATGTGCATCAAGAGGCTGCA 434  
|||||  
101 ValProIleProLeuGlySerLysAsnMetCysIleLysArgLeuG1 117  
  
435 GATTAAACCCAGAACCTTTAGTGGACTCATTATTTAAATCCCTTTACC 484  
|||||  
117 nleLysProArgSerPheSerGlyLeuThrTyrLeuLysSerLeuTyrL 134  
  
485 TGGATGGAACACAGCTTACTAGAGATACCGAGGCGCTCCCGCTAGCTTTA 534  
|||||  
134 euaspGlyasnGlnLeuLeuGluIleProGlnGlyLeuProProSerLeu 150  
  
535 CAGCTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAGAGAA 584  
|||||  
151 GlnLeuLeuSerLeuGluAlaasnAsnIlePheSerIleArgLysGluAs 167  
  
585 TCTAACAGACTGGCCCAACATAGAAATACTCTACCTGGGCCCAAACTGTT 634  
|||||  
167 nLeuThrGluLeuAlaAsnIleGluIleLeuTyrLeuGlyGlnAsnCysT 184  
  
635 ATTATCGAAATCCTTCTTATGTTTCATATTCAATAGAGAAAGATGCTTC 684  
|||||  
184 yrTyrtargasnProCysTyrValSerTyrSerIleGluLysaspAlaPhe 200  
  
685 CTAAACTTGACAAAGTTAAAGTGTCTCTCCCTGAAAGATAACAATGTAC 734

|||||  
201 LeuAsnLeuThrLysLeuLysValLeuSerLeuLysAspAsnValTh 217  
735 AGCGTCCCTACTGTTTGGCATCTACTTTAACAGAACTATATCTCTACA 784  
217 rAlaValProThrValLeuProSerThrLeuThrGluLeuTyLeuTyra 234  
785 ACAACATGATTGCAAAATCCAGAAGATGATTTTAATAACCTCAACCAA 834  
234 snAsnMetIleAlaLysIleGlnGluAspPheAsnAsnLeuAsnGln 250  
835 TTACAAATCTTGACCTAAAGTGAATTTGCCCTCGTTGTTTATATGCCCC 884  
251 LeuGlnIleLeuAspLeuSerGlyAsnCysProArgCysTyraAsnAlaPr 267  
885 ATTTCTTGTGCGCGCTGTAATAATAATTTCTCCCTACAGATCCCTGTAA 934  
267 oPheProCysAlaProCysLysAsnAsnSerProLeuGlnIleProVala 284  
935 ATGCTTTTGTATGCGCTCACAGAATTTAAAGTTTACGCTACACAGTAAC 984  
284 snAlaPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300  
985 TCCTTTCAGCATGCCCCCAAGATGGTTTAAAGAACATCAACAACTCCA 1034  
301 SerLeuGlnHisValProProArgTrrPheLysAsnIleAsnLysLeuGl 317  
1035 GGAAGTGGATCTGTCACAAATCTTGGCCAAAGAAATTTGGGATGCTA 1084  
317 nGluLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaL 334  
1085 AATTTCGCAATTTCTCCCGCTCATCAATTTGGATCTGCTTCAAT 1134  
334 ysPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350  
1135 TTGGAATTCAGGTCTATCGTCATCTATGAATCTATCAAGCATTTTC 1184  
351 PheGluLeuGlnValTyArgAlaSerMetAsnLeuSerGlnAlaPheSe 367  
1185 TTCACTGAAAGCTGAAATTTGCGGATCAGAGGATATGCTTTAAG 1234  
367 rSerLeuLysSerLeuLysIleLeuArgIleArGGlyTyraValPheLysG 384  
1235 AGTTGAAAGCTTTAACTCTCGCATTTACATAATCTTCAAAATCTTGAA 1284  
384 luLeuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400  
1285 GTTCTTGATCTTGGCACTAACTTTATAAAATTTGCTAACCTCAGCATGTT 1334  
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417  
1335 TAAACAATTTAAAGACTGAAAGTCATAGATCTTTCAGTGAATAAATAT 1384  
417 eLysGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIleS 434  
1385 CACCTTCAGGAGATTCAAAGTGAAGTTGGCTTCTGCTCAAAATGCCAAGT 1434  
434 erProSerGlyAspSerSerGluValGlyPheCysSerAsnAlaArgThr 450  
1435 TCTGTAGAAAGTTATGAACCCCGGTCCTGGAACAAATTCATATTTCAG 1484  
451 SerValGluSerTyraGluProGlnValLeuGluGlnLeuHisTyraPheAr 467  
1485 ATATGATAAGTATGCAAGGAGTTGCAGATTCAAAACAAAGAGCTCTTT 1534  
467 gTyraSpLysTyraAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484  
1535 TCATGCTCTGTTAATGAAAGCTGCTACAAGTATGGGAGACCTTTGGATCTA 1584  
484 heMetSerValAsnGluSerCysTyrrLysTyrrGlyGlnThrLeuAspLeu 500  
1585 AGTAAATAGTATATTTTTCAGTCCCTGATTTTCAGCATCTTTTC 1634  
|||||

501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517  
1635 TTTCCCTCAAATGCTGAATCTGTACAGAAATCTCATTAGCCAAACTCTTA 1684  
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534  
1685 ATGGCAGTGAATTTCAACCTTTAGCAGAGCTGAGATATTTGGACTTCTCC 1734  
534 snGlySerGluPheGlnProLeuAlaGluLeuArgTyraLeuAspPheSer 550  
1735 AACAAACCGGCTTGATTTACTTCCATTCAACAGCATTTGAAGAGCTTCACAA 1784  
551 AsnAsnArgLeuAspLeuHisSerThrAlaPheGluGluLeuHisLys 567  
1785 ACTGGAAGTTCTGGATATATAGCAGTATAGCCATTTTTCATCATCAGAAG 1834  
567 sLeuGluValLeuAspIleSerSerAsnSerHisTyraPheGlnSerGluG 584  
1835 GAATTTACTCATATGCTTAACTTTTACCAAGAACCTTAAAGTTTCTGCAGAAA 1884  
584 lyleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600  
1885 CTGATGATGAACGACAAATGACATCTCTCTCCACCAGCAGGACCATGGA 1934  
601 LeuMetMetAsnAspAsnAspIleSerSerSerThrSerArgThrMetGl 617  
1935 GAGTGTGCTCTTTAGAACTCTGGAATTCAGAGGAATTCACATTAGATGTTT 1984  
617 uSerGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspValL 634  
1985 TATGGAGAGAAGGTGATTAACAGATACTTACAATTTATTCAGAAATCTGCTA 2034  
634 euTrpArgGluGlyAspAsnArgTyraLeuGlnLeuPheLysAsnLeuLeu 650  
2035 AAATTAGAGGAATTAGACATCTCTAAAATTTCCCTAAGTTTCTTGCTTC 2084  
651 LysLeuGluGluLeuAspIleSerLysAsnSerLeuSerPheLeuProSe 667  
2085 TGCAGTTTTTGTGATGGTATGCCTCCAAATCTTAAAGAAATCTCTCTTTGGCCA 2134  
667 rGlyValPheAspGlyMetProProAsnLeuLysAsnLeuSerLeuAlaL 684  
2135 AAAATGGCTCAAATCTTTTCAGTTGGAAGAACTCCAGTGTCTTAAAGAAC 2184  
684 ysAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
2185 CTGGAACCTTTGGACCTCAGCCACAACTGACCACCTGTCCTCCTGAGAG 2234  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrThrValProGluAr 717  
2235 ATTATCCAAGTGTTCAGAGCCTCAAGATCTGATTCATTCTTAAGAATAATC 2284  
717 gLeuSerAsnCysSerArgSerLeuLysAsnLeuIleLeuLysAsnAsnG 734  
2285 AAATCAGGAGTCTGAGAAAGTATTTTCTACAAGATGCTTCCAGTTGCGA 2334  
734 InIleArgSerLeuThrLysTyraPheLeuGlnAspAlaPheGlnLeuArg 750  
2335 TATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAGACCACTT 2384  
751 TyraLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
2385 CCCAGAAAATGCTCTCAACAACTGAGATGTTGCTTTTGCATCATTAATC 2434  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuLeuHisHisAsnA 784  
2435 GGTTTCTGTGCACCTGTGATGCTGCTGCTGTTTCTGTTGGTGGTAAACCAT 2484  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpValAsnHis 800  
2485 ACGGAGGTGACTATCTTACCTGGCCACAGATGTGACTTGTCTGGGGCC 2534  
801 ThrGluValThrIleProTyraLeuAlaThrAspValThrCysValGlyPr 817  
|||||

2535 AGGAGCACAAAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
|||||  
817 oGlyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyrThrCysG 834  
2585 AGTTAGATCTGACTAACTGATGTTCTGTTCTACATTTCCATATCTGTATCT 2634  
|||||  
834 luleAspLeuThrAsnLeuIleLeuPheSerLeuSerIleSerValSer 850  
2635 CTCTTTCTCTGATGATGATGACAGCAAGTCACTCTATTCTTGGGATGT 2684  
|||||  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyrPheTrpAspVa 867  
2685 GTGGTATATTTACATTTCTGTAAAGCCCAAGATAAAGGGGTATCAGCGTC 2734  
|||||  
867 lTrpTyrIleTyrHisPheCysLysAlaLysIleLysGlyTyrGlnArgL 884  
2735 TAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTATGATGACACTAA 2784  
|||||  
884 eulleSerProAspCysCysTyrAspAlaPheIleValTyrAspThrLys 900  
2785 GACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGTGGCCAACTGGA 2834  
|||||  
901 AspProAlaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuGI 917  
2835 AGACCCCAAGAGAGAAACATTTTAAATTTATGTTCTCGAGGAAAGGACTGGT 2884  
|||||  
917 uAspProArgGluLysHisPheAsnLeuCysLeuGluGluArgAspTrpL 934  
2885 TACCAGGCGCAGCTCTGGAAGAACTTTCCAGAGCATACAGCTTAGC 2934  
|||||  
934 euProGlyGlnProValLeuGluAsnLeuSerGlnSerIleGlnLeuSer 950  
2935 AAAAGACAGCTGTTGTGATGACAGACAAGTATGCAAGACTGAAATTT 2984  
|||||  
951 LysLysThrValPheValMetThrAspLysTyrAlaLysThrGluAsnPh 967  
2985 TAGATAGCATTTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTG 3034  
|||||  
967 eLysIleAlaPheTyrLeuSerHisGlnArgLeuMetAspGluLysValA 984  
3035 ATCTGATATCTTGCATATTTCTTGAAGCCCTTTCAGAGCTCAAGTTC 3084  
|||||  
984 spValIleIleLeuIlePheLeuGluLysProPheGlnLysSerLysPhe 1000  
3085 CTCAGCTCCGGAAAGGCTCTGTGGAGTCTGTCTCTTGCAGTGGCCAAC 3134  
|||||  
1001 LeuGlnLeuArgLysArgLeuCysGlySerSerValLeuGluTrpProTh 1017  
3135 AAACCCGCAAGCTCACCCATCTTCTGGCAGTGTCTAAAGAACGCCCTGG 3184  
|||||  
1017 rAsnProGlnAlaHisProTyrPheTrpGlnCysLeuLysAsnAlaLeuA 1034  
3185 CCACAGACAATCATGTGGCTGTAGTCAGGTGTCAGGGAACGGTC 3231  
|||||  
1034 laThrAspAsnHisValAlaTyrSerGlnValPheLysGluThrVal 1049

seq\_name: /SDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAB44323

seq\_documentation\_block:

ID: AAB44323 standard; Protein; 1049 AA.

XX AC: AAB44323;

XX DT: 08-FEB-2001 (first entry)

XX DE: Human PRO285 protein sequence SEQ ID NO:496.

XX KW: Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;

XX KW: expressed sequence tag; detection; cancer.

XX OS: Homo sapiens.

XX

PN WO200053756-A2.

XX 14-SEP-2000.

XX 18-FEB-2000; 2000WO-US04341.

XX 08-MAR-1999; 99WO-US05028.

XX 12-MAR-1999; 99US-0123957.

XX 29-MAR-1999; 99US-0126773.

XX 21-APR-1999; 99US-0130232.

XX 28-APR-1999; 99US-0131445.

XX 14-MAY-1999; 99US-0134287.

XX 23-JUN-1999; 99US-0141037.

XX 26-JUL-1999; 99US-0145698.

XX 29-OCT-1999; 99US-0162506.

XX 30-NOV-1999; 99WO-US28313.

XX 02-DEC-1999; 99WO-US28551.

XX 02-DEC-1999; 99WO-US28565.

XX 16-DEC-1999; 99WO-US30095.

XX 30-DEC-1999; 99WO-US31243.

XX 30-DEC-1999; 99WO-US31274.

XX 05-JAN-2000; 2000WO-US00219.

XX 06-JAN-2000; 2000WO-US00277.

XX 06-JAN-2000; 2000WO-US00376.

XX (GETH ) GENENTECH INC.

PA Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;

XX Ferrara N, Fillvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;

PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;

PI Kljavin LJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;

PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;

XX WPI: 2000-611443/58.

DR N-PSDB; AAC78583.

XX Novel PRO polypeptides and polynucleotides used in detection methods,

PT to target bioactive molecules to specific cells, and to modulate

PT cellular activities -

XX Claim 12; Fig 209; 636pp; English.

XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed

CC sequence tag) sequences which encode secreted or transmembrane PRO

CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic

CC activity. The polynucleotides and polypeptides can be used for detecting

CC the presence of PRO polypeptides in samples for linking bioactive

CC molecules to cells and for modulating biological activities of cells,

CC using the polypeptides for specific targeting. The polypeptide targeting

CC can be used to kill the target cells, e.g. for the treatment of cancers.

CC The polypeptide pairs provide specific targeting of bioactive molecules

CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in

CC the isolation of the PRO polynucleotide sequences.

XX SQ Sequence 1049 AA;

alignment\_scores:

Quality: 1049.00 Length: 1049

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAB44323 ..

Align seg 1/1 to: AAB44323 from: 1 to: 1049

85 ATGCTGTTTCCAAATGTGGACACTGAAGACAAATTTCTTATCTCTTTTAA 134

|||||

1 MetValPheProMetTrpThrLeuLysArgGlnIleLeuIleLeuPheAs 17

|||||

135 CATATCTCTTAATTTCCAAACTCTCTGGGGCTAGATGTTTCTTAAACTC 184

|||||

|||||

17 nIleIleLeuIleSerLysLeuLeuGlyAlaArgTrpPheProLysThrL 34  
185 TGCCCTGTGATGTCACCTCTCGATGTTCCAAAGAACCATGTGATCGTGGAC 234  
34 euProCysAspValThrLeuAspValProLysAsnHisValIleValAsp 50  
235 TGCACAGACAAAGCATTTGACAGAAATTCCTCGAGGTATTCCCCACGAAAC 284  
51 CysThrAspLysHisLeuThrGluIleProGlyGlyIleProThrAsnTh 67  
285 CAGAACCTCACCTCACCTCACCATTAACACATACCAGACATCTCCCGAGCGT 334  
67 rThrAsnLeuThrLeuThrIleAsnHisIleProAspIleSerProAlas 84  
335 CCATTTCACAGACTGGACCATCTCGTAGAGATCGATTTCAGATCCAACTGT 384  
84 erPheHisArgLeuAspHisLeuValGluIleAspPheArgCysAsnCys 100  
385 GTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCA 434  
101 ValProIleProLeuGlySerLysAsnMetCysIleIysArgLeuGI 117  
435 GATTAAACCCAGAGCTTTAGTGGACTCACTTATTTAAATCCCTTTACC 484  
117 nIleLysProArgSerPheSerGlyLeuThrTyrLeuLysSerLeuTyrL 134  
485 TGGATGGAACACAGCTACTAGAGATACCGCAGGCGCTCCCGCTAGCTTA 534  
134 euAspGlyAsnGlnLeuLeuGluIleProGlnGlyLeuProProSerLeu 150  
535 CAGCTTCTCAGCTTGGGCAACAACATCTTTTCCATCAGAAAGAGAA 584  
151 GlnLeuLeuSerLeuGluAlaAsnAsnIlePheSerIleArgLysGluAs 167  
585 TCTAACAGAACTGGCAACATAGAAATACCTCTACCTGGGCCAAACTGTT 634  
167 nLeuThrGluLeuAlaAsnIleGluIleLeuTyrLeuGlyGlnAsnCyst 184  
635 ATTATCGAAATCCTGTTATGTTTCATATCAATAGAGAAAGATCGCTTC 684  
184 yrTyrArgAsnProCysTyrValSerTyrSerIleGluLysAspAlaPhe 200  
685 CTAAACTTGACAAAGTTAAAGTCCTCCCTGAAAGATAACAAATGTCA 734  
201 LeuAsnLeuThrLysLeuLysValLeuSerLeuLysAspAsnValTh 217  
735 AGCGTCCCTACTGTTTGGCATCTACTTTAACAGAACTATATCTCTACA 784  
217 rAlaValProThrValLeuProSerThrLeuThrGluLeuTyrLeuTyrA 234  
785 ACAACATGATTGCAAAATCCAAGAAGATGATTTTAATACCTCAACCAA 834  
234 snAsnMetIleAlaLysIleGlnGluAspAspPheAsnAsnLeuAsnGln 250  
835 TTACAAATCTTGACCTAAGTGAAATTCGCCCTCGTTGTTATATGCCCC 884  
251 LeuGlnIleLeuAspLeuSerGlyAsnCysProArgCysTyrAsnAlaPr 267  
885 ATTTCTTGTGCCCGCTGTAATAATTAATCTCCCTCCCTACAGATCCCTGTAA 934  
267 oPheProCysAlaProCysLysAsnSerProLeuGlnIleProVala 284  
935 ATGCTTTTGTGCGCTGACAGAAATTAAGTTTACGTCTACACAGCTAAC 984  
284 snAlaPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300  
985 TCTCTTACAGATGTCGCCCAAGATGGTTTAAAGACATCAACAACTCCA 1034  
301 SerLeuGlnHisValProProArgTrpPheLysAsnIleAsnLysLeuGI 317  
1035 GGAACGTGATCTGTCCCAAACTTCTTGCCCAAGAAATTTGGGGATGCTTA 1084  
317 nGluLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaL 334

1085 AATTTCTGCATTTTCTCCAGCTCATCCAATTTGGATGTGTCTTCAAT 1134  
334 ysPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350  
1135 TTTGAACTTCAGGTCTATCGTGCATCTATGAATCTATCACAGCATTTTC 1184  
351 PheGluLeuGlnValTyrArgAlaSerMetAsnLeuSerGlnAlaPheSe 367  
1185 TTCACTGAAAGCCTGAAAATTTCTCGGGATCAGAGGATATGCTTTAAAG 1234  
367 rSerLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysG 384  
1235 AGTTGAAAAGCTTTAACTCTCGCCATTACATAATCTTCAAAATCTTGAA 1284  
384 luleuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400  
1285 GTTCTTGATCTTGGCACTAACTTTATAAANTTGTAACTCAGCATGTT 1334  
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417  
1335 TAAACAATTTAAAGACTGAAAGTCATAGATCTTTTCAGTGAATAAATAT 1384  
417 eLysGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIles 434  
1385 CACCTTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAATGCCAGAACT 1434  
434 erProSerGlyaspSerSerGluValGlyPheCysSerAsnAlaArgThr 450  
1435 TCTGTAGAAAGTTATGAACCCAGGTCCTGGACAAATTAATATTTCAG 1484  
451 SerValGluSerTyrGluProGlnValLeuGluGlnLeuHisTyrPheAr 467  
1485 ATATGATAAGTATGCAAGGAGTTGCAGATTCAAAACAAAGAGGCTTCTT 1534  
467 gTyrAspLysTyrAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484  
1535 TCATGTCGTGTTAATGAAAGCTGCTACAAGTATGGCGCAGACCTTGGATCTA 1584  
484 heMetSerValAsnGluSerCysTyrGlyGlnThrLeuAspLeu 500  
1585 AGTAAATAAGTATATATTTTGTCAAGTCCTCTGATTTTCAGCATCTTTC 1634  
501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517  
1635 TTTCTCAAATGCCTGAATCTGTACGGAATCTCATTTAGCCAAACTCTTA 1684  
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534  
1685 ATGGCAGTGAATTCACACCTTTAGCAGAGCTGAGATATTGGACTTCTCC 1734  
534 snGlySerGluPheGlnProLeuAlaGluLeuArgTyrLeuAspPheSer 550  
1735 AACAAACCGGCTTGATTTACTCCATTCAACAGCATTTTGAAGAGCTTCACAA 1784  
551 AsnAsnArgLeuAspLeuLeuHisSerThrAlaPheGluGluLeuHisLy 567  
1785 ACTGGAAGTTCGGGATATAAGCAGTAATAGCCATTTATTTCAATCAGAAG 1834  
567 sLeuGluValLeuAspIleSerSerAsnSerHisTyrPheGlnSerGluG 584  
1835 GAATTACTCATATGCTAAACTTTTACCAGAACCTTAAAGGTTCTGCGAGAA 1884  
584 lylleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600  
1885 CTGATGATGAACGACAAATGACATCTCTTCTCCACGACGAGGACCATGGA 1934  
601 LeuMetMetAsnAspAsnAspIleSerSerSerThrSerArgThrMetGI 617  
1935 GAGTGAATCTCTAGAACTCTGGAATTCAGAGAAATCACTTAGATGTTT 1984  
617 uSerGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspValL 634

1985 TATGGAGAGAGGTGATACAGATCTTACAATATTCAAGAAATCTGCTA 2034  
|||||  
634 euTrpArgGluGlyAspAsnArgTyrLeuGlnLeuPheLysAsnLeuLeu 650  
2035 AAATTAGAGGAATTAGACATCTCTAAAATCCCTAAAGTTCTTCGCTTC 2084  
|||||  
651 LysLeuGluGluLeuAspLysSerLysAsnSerLeuSerPheLeuProse 667  
2085 TGGAGTATTTGATGGTATGCCCTCAAAATCTAAAGAAATCTCTTTGGCCA 2134  
667 rGlyValPheAspGlyMetProProAsnLeuLysAsnLeuSerLeuAlaL 684  
2135 AAAATGGGCTCAAACTTTCACTGTTGAAGAACTCCAGTGTCTAAAGAAC 2184  
|||||  
684 ysAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
2185 CTGGAACTTTGGACCTGACCCACACACCACTGACCACCTGCTCCCTGAGAG 2234  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrThrValProGluAr 717  
2235 ATTATCCAACCTGTTCCAGAAGCCTCAAGAATCTGATTCTTTAAAGAATAATC 2284  
717 gLeuSerAsnCysSerArgSerLeuLysAsnLeuLleLeuLysAsnAsnG 734  
2285 AAATCAGGAGTCTGAGGAAGTATTTCTTACAAGATGCCCTCCAGTTGGCA 2334  
734 InIleArgSerLeuThrLysTyrPheLeuGlnAspAlaPheGlnLeuArg 750  
2335 TATCTGGATCTCAGCTCAATAAATCCAGATCATCCAAAGACCAGCTT 2384  
|||||  
751 TyrLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
2385 CCAGAAAAATGCTCTCAACAATCTCAAGATGTTGCTTTTGCATCAATAATC 2434  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuLeuHisHisAsnA 784  
2435 GGTTCCTGTCACCTGTGATGCTGTGCTGTTGCTGTGGGTAAACCAT 2484  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpTrpValAsnHis 800  
2485 ACGGAGTCACTATCTCTTACCTGGCCACAGATGCTACTGTCTGGGGCC 2534  
801 ThrGluValThrIleProTyrLeuAlaThrAspValThrCysValGlyPr 817  
2535 AGGACACACAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
817 oGlyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyrThrCysG 834  
2585 AGTTAGATCTGACTAACCTGATCTCTTCTCCTACCTTTCCATATCTGTATCT 2634  
834 luLeuAspLeuThrAsnLeuIleLeuPheSerLeuSerIleSerValSer 850  
2635 CTCTTTCTCATGCTGATGATGACAGCAAGTCACTCTATTCTGGGATCT 2684  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyrPheTrpAspVa 867  
2685 GTGGTATATTTACCATTTCTGTAAAGCCCAAGATAAAGGGGTATCAGCGTC 2734  
867 lTrpTyrIleTyrHisPheCysLysAlaLysIleLysGlyTyrGlnArgL 884  
2735 TAATATCACAGACTGTTGTATGATGCTTTTATTTATGTTATGATGACACTAA 2784  
884 euIleSerProAspCysCysTyrAspAlaPheIleValTyrAspThrLys 900  
2785 GACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGGTGGCCAACTGGA 2834  
901 AspProAlaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuGl 917  
2835 AGACCCAAAGAGAGAAACATTTTAAATTTATCTCTCGAGGAAAGGACTGGT 2884  
917 uAspProArgGluLysHisPheAsnLeuCysLeuGluGluArgAspTrpL 934  
2885 TACCAGGGCAGCCAGTCTCTGGAAAAACCTTTTCCAGAGCATACAGCTTAGC 2934

|||||  
934 euProGlyGlnProValLeuGluAsnLeuSerGlnSerIleGlnLeuSer 950  
2935 AAAAAGACAGTCTTTGTGTATGACAGACAAGTATGCAAGACTGAAAATTT 2984  
951 LysLysThrValPheValMetThrAspLysTyrAlaLysThrGluAsnPh 967  
2985 TAAGTAGACATTTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTG 3034  
967 elysIleAlaPheTyrLeuSerHisGlnArgLeuMetAspGluLysValA 984  
3035 ATGTGATATATCTTGATATTTCTTGAGAACCCCTTTCAGAAAGTCCAAGTTC 3084  
984 spValIleIleLeuIlePheLeuGluLysProPheGlnLysSerLysPhe 1000  
3085 CTCCAGCTCCGGAAAGGCTCTGTGGAGTCTGTCTTGTAGTGGCCAAC 3134  
1001 LeuGlnLeuArgLysArgLeuCysGlySerSerValLeuGluTrpProth 1017  
3135 AAACCCGCAAGCTCACCCATCTTCTGCGAGTGTCTTAAAGAACGCCCTGG 3184  
1017 rAsnProGlnAlaHisProTyrPheTrpGlnCysLeuLysAsnAlaLeuA 1034  
3185 CCACAGACAATCATGTGGCCTATAGTCAGTGTTCAGGAAACCGTTC 3231  
1034 laThrAspAsnHisValAlaTyrSerGlnValPheLysGluThrVal 1049

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAU12350

seq\_documentation\_block:

ID AAU12350 standard; Protein; 1049 AA.

XX AAU12350;

DT 24-OCT-2001 (first entry)

XX Human PRO285 polypeptide sequence.

Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
adipocyte; A-peptide; factor VIIA; gene therapy.

OS Homo sapiens.

XX WO200140466-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US32678.

PR 01-DEC-1999; 99WO-US28301.

PR 01-DEC-1999; 99WO-US28634.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28564.

PR 02-DEC-1999; 99WO-US28565.

PR 16-DEC-1999; 99US-0170262.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 20-DEC-1999; 99WO-US30999.

PR 30-DEC-1999; 99WO-US31243.

PR 06-JAN-2000; 2000WO-US00277.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 18-FEB-2000; 2000WO-US04342.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 01-MAR-2000; 2000WO-US05004.

PR 21-MAR-2000; 2000WO-US05601.

PR 20-MAR-2000; 2000WO-US07377.

PR 21-MAR-2000; 2000WO-US07532.

PR 30-MAR-2000; 2000WO-US08439.



PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14041.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US10873.  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart RA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR N-PSDB: AAS21422.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX  
 PS Claim 12; Fig 358; 813pp; English.  
 XX  
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 SQ Sequence 1049 AA;

alignment\_scores:  
 Quality: 1049.00 Length: 1049  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAU12350

Align seg 1/1 to: AAU12350 from: 1 to: 1049

85 ATGGTGTTCCTCAATGTGGACACTGAAGACACAAATCTTATCTTTTAA 134  
 1 MetValPheProMetThrPheLeuLysArgGlnLeuLeuPheAsn 17  
 135 CATATCTCAATTTCAAACCTCTGGGGCTAGATGGTTTCTCTAAACTC 184  
 17 nleuLeuLeuLeuSerLysLeuLeuGlyAlaArgTrpPheProLysThrL 34  
 185 TGGCTGTGATGTCACTCTGGATGTTCCTCAAGAACCATGTGCTGGAC 234  
 34 euProCysAspValThrLeuAspValProLysAsnHisValIleValAsp 50  
 235 TGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCTCCACGAACAC 284  
 51 CysThrAspLysHisLeuThrGluIleProGlyGlyIleProThrAsnTh 67  
 285 CAGGAACCTCACCTCACCATTAAACACATACACAGACATCTCCCGAGCCT 334

67 rThrAsnLeuThrLeuThrIleAsnHisIleProAspIleSerProAlas 84  
 335 CCTTTTCACAGACTGGACCATCTGGTAGAGATCGATTTTCAGATGCAACTGT 384  
 84 erPheHisArgLeuAspHisLeuValGluIleAspPheArgCysAsnCys 100  
 385 GTACCTATTCCACTGGGGTCAAAAAACAACATGTGTCATCAAGAGGCTGCA 434  
 101 ValProIleProLeuGlySerLysAsnMetCysIleLysArgLeuGlu 117  
 435 GATTAAACCCAGAGCTTTAGTGGACTCACTTATTTAAATCCCTTTACC 484  
 117 nleLysProArgSerPheSerGlyLeuThrTyrLeuLysSerLeuTyrL 134  
 485 TGGATGGAAACAGCTACTAGATACCCGAGGCTCCCGCTAGCTTTA 534  
 134 euAspGlyAsnGlnLeuLeuGluIleProGlnGlyLeuProSerLeu 150  
 535 CAGCTTCTCAGCTTGAGGCCAACACATCTTTCCATCAGAAAGAGAA 584  
 151 GlnLeuLeuSerLeuGluAlaAsnAsnIlePheSerIleArgLysGluAs 167  
 585 TCTAACAGAACTGGCCAACATAGAAATCTACTCTGGGCCAAACTGTT 634  
 167 nLeuThrGluLeuAlaAsnIleGluIleLeuTyrLeuGlyGlnAsnCysT 184  
 635 ATTATCGAAATCCTTGTATTATGTTTCATATTCATAGAGAAAGATGCCCTC 584  
 184 yTyrArgAsnProCysTyrValSerTyrSerIleGluLysAspAlaPhe 200  
 685 CTAAACTTCACAAAGTTAAAGTCTCCCTGAAAGATAACAATGTCA 734  
 201 LeuAsnLeuThrLysLeuLysValLeuSerLeuLysAspAsnValTh 217  
 735 AGCGTCCCTACTGTTTCCCATCTACTTTAACAGAACTATATCTCTACA 784  
 217 rAlaValProThrValLeuProSerThrLeuThrGluLeuTyrLeuTyrA 234  
 785 ACACATGATTCCAAAATCCAGAGATGATTTTAATAACCTCAACCAA 834  
 234 snAsnMetIleAlaLysIleGlnGluAspAspPheAsnAsnLeuAsnGln 250  
 835 TTACAAATTTCTGACCTAAGTGGAAATTCCTCGTTGTATAATGCCCC 884  
 251 LeuGlnIleLeuAspLeuSerGlyAsnCysProArgCysTyrAsnAlaPr 267  
 885 ATTCTCTGTGCGCGTGTAAAATAATTTCTCCCTACAGATCCCTGTAA 934  
 267 oPheProCysAlaProCysLysAsnAsnSerProLeuGlnIleProValA 284  
 935 ATGCTTTTTCATCGCTGACAGAAATTTAAAGTTTACGTCTACACAGTAAC 984  
 284 snAlaPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300  
 985 TCTCTTCAGCATGTGCCCCCAAGATGGTTTAAAGAACATCAACAACCTCCA 1034  
 301 SerLeuGlnHisValProProArgTrpPheLysAsnIleAsnLysLeuGlu 317  
 1035 GGAATCGGATCTGCTCCCAAACTTCTTGCCCAAGAAATTTGGGGATGCTA 1084  
 317 nGluLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaL 334  
 1085 AATTTCTGCATTTCTCCCGACCTCATCTCAATTCGATCTGCTTTCAAT 1134  
 334 ysPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350  
 1135 TTTGAACCTCAGGTCTATCGTCATCTCTGAATCTCTCAAGACATTTTC 1184  
 351 PheGluLeuGlnValTyrArgAlaSerMetAsnLeuSerGlnAlaPheSe 367  
 1185 TTCCTGAAAAGCCTGAAAATTTCTGGGGATCAGAGGATATGCTTTTAAAG 1234



367 rSerLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysG 384  
1235 AGTTGAAAGCTTTAACTCGCGCATTAACATCTTCAAAATCTTCAA 1284  
|||||  
384 luLeuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400  
1285 GTTCTTGATCTGGCAGCTAACTTTATAAAATTTGCTAACCTCAGCATGTT 1334  
|||||  
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417  
1335 TAAACAATTTAAAGACGAAAGTCATAGATCTTTCAGTGAATAAAATAT 1384  
|||||  
417 eLysGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIleS 434  
1385 CACCTTCAGGAGATCAAGTGAAGTTGGCTCTGCTCAAATGCCAGAACT 1434  
|||||  
434 erProSerGlyAspSerSerGluValGlyPheCysSerAsnAlaArgThr 450  
1435 TCTGTAGAAAGTTATGAACCCAGGCTCTCGGAACAATTTACATTTATTCAG 1484  
|||||  
451 SerValGluSerTyrGluProGlnValLeuGluGlnLeuHisTyrPheAr 467  
1485 ATATGATAAGTATGCAAGGAGTTGCAGATTCAAAACAAAGAGGCTTCTT 1534  
|||||  
467 gTyrAspLysTyrAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484  
1535 TCATCTCTCTTAATGAAAGCTGCTACAAGTATGGCAGACCTTTGGATCTA 1584  
|||||  
484 heMetSerValAsnGluSerCysTyrLysTyrGlyGlnThrLeuAspLeu 500  
1585 AGTAAAAATAGTATATTTTTGTGCAAGTCTCTGATTTTCAGCATCTTTC 1634  
|||||  
501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517  
1635 TTTTCTCAAAATCCCTGAATCTGTCAGGAATCTCATTAGCCCAACTCTTA 1684  
|||||  
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534  
1685 ATGGCAGTGAATTCACACTTTAGCAGAGCTGAGATATTTGGACTTCTCC 1734  
|||||  
534 snGlySerGluPheGlnProLeuAlaGluLeuArgTyrLeuAspPheSer 550  
1735 ACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAA 1784  
|||||  
551 AsnAsnArgLeuAspLeuLeuHisSerThrAlaPheGluGluLeuHisLy 567  
1785 ACTGGAAGTCTGGATATAAGCAGTAATAGCCATTATTTCAATCAGAAG 1834  
|||||  
567 sleuGluValLeuAspIleSerSerAsnSerHisTyrPheGlnSerGluG 584  
1835 GAATTAATCATATGCTAAACTTTACCAAGAACTAAAGGTTCTGCAGAAA 1884  
|||||  
584 lyIleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600  
1885 CTGATGATGAACGACAATGACATCTCTTCCCTCCACGACGAGCAGCATGA 1934  
|||||  
601 LeuMetMetAsnAspAsnAspIleSerSerSerThrArgThrMetGl 617  
1935 GAGTCAGTCTCTAGAACTCTGGAATTCAGAGGAATCAGTATAGATGTTT 1984  
|||||  
617 uSerGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspValL 634  
1985 TATGGAGAGAGGCTGATACAGATACCTTACAATTTATCAAGAAATCTGCTA 2034  
|||||  
634 euTrpArgGluGlyAspAsnArgTyrLeuGlnLeuPheLysAsnLeuLeu 650  
2035 AAATTAGAGGAATTAGACATCTCTAAAATTCCTTAAGTTTCTGCGCTTC 2084  
|||||  
651 LysLeuGluGluLeuAspIleSerLysAsnSerLeuSerPheLeuProSe 667  
2085 TCGAGCTTTTGTGATGCTATGCTCCCTCAAAATCTAAAGAAATCTCTCTTTGGCCA 2134  
|||||  
667 rGlyValPheAspGlyMetProProAsnLeuLysAsnLeuSerLeuAlaL 684

2135 AAAATGGGCTCAAAATCTTTTCAGTTTGGAAAGAACTCCAGTGTCTAAAGAAC 2184  
|||||  
684 ysAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
2185 CTGGAACACTTTGGACCTCAGCCACACCACTGACCACTGCTCCCTGAGAG 2234  
|||||  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrThrValProGluAr 717  
2235 ATTTATCCAACCTGTTCCAGAACCTCAAGAACTGTGATCTTAAAGATAATC 2284  
|||||  
717 gluSerAsnCysSerArgSerLeuLysAsnLeuIleLeuLysAsnAsnG 734  
2285 AAATCAGGAGCTGACGAAAGTATTTTCTACAAGATGCTTCCAGTTGGCA 2334  
|||||  
734 lnIleArgSerLeuThrLysTyrPheLeuGlnAspAlaPheGlnLeuArg 750  
2335 TATCTGGATCTCAGCTCAAAATAAAATCCAGATGATCCAAAAGACCAGCTT 2384  
|||||  
751 TyrLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
2385 CCCAGAAAATGCTCTCAACAATCTGAAGATGTTGCTTTTGCATCATATATC 2434  
|||||  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuHisHisAsnA 784  
2435 GGTFTTCTGTCACCTGTCATGCTGTGTTGCTGTGTTGCTGTGTTAACCAT 2484  
|||||  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpValAsnHis 800  
2485 ACGGAGGTGACTATTCTCTTACTGGCCACAGATGTGACTTGTGTGGGGCC 2534  
|||||  
801 ThrGluValThrIleProTyrLeuAlaThrAspValThrCysValGlyPr 817  
2535 AGGACACACAAAGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
|||||  
817 oclyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyrThrCysG 834  
2585 AGTTAGATCTGACTAACCTGATTTCTGTCTCACTTTCCATATCTGTATCT 2634  
|||||  
834 luLeuAspLeuThrAsnLeuIleLeuPheSerLeuSerIleSerValSer 850  
2635 CTCCTTCTCATGGTGATGATGACACAGATCACCTCTATTCTTGGGATGT 2684  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyrPheTrpAspVa 867  
2685 GTGGTATATTTACCATTCTGTAAAGCCCAAGATAAAGGGTATCAGCGTC 2734  
|||||  
867 lTrpTyrIleTyrHisPheCysLysAlaLysIleLysGlyTyrGlnArgL 884  
2735 TAATATCACAGACTGTTGCTATGATGCTTTTATTGTGTATGACACATAA 2784  
|||||  
884 euIleSerProAspCysCysTyrAspAlaPheIleValTyrAspThrLys 900  
2785 GACCCAGCTGTCAGCCAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGA 2834  
901 AspProAlaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuGl 917  
2835 AGACCCAAAGAGAGAAACATTTTAAATTTATGCTCGAGGAAAGGACTGCT 2884  
|||||  
917 uAspProArgGluLysHisPheAsnLeuCysLeuGluArgAspTrpL 934  
2885 TACCCAGGCGCAGCTTCTTGGAAAACTTTCCAGAGCATACAGCTTAGC 2934  
|||||  
934 euProGlyGlnProValLeuGluAsnLeuSerGlnSerIleGlnLeuSer 950  
2935 AAAAAGACAGCTTTGCTGATGACACAGATGATGCAAGACTGAAAATTT 2984  
951 LysLysThrValPheValMetThrAspLysTyrAlaLysThrGluAsnPh 967  
2985 TAAAGATACATTTTACTTGTCTCCCATCAGAGGCTCATGGATGAAAAGTTG 3034  
|||||  
967 eLysIleAlaPheTyrLeuSerHisGlnArgLeuMetAspGluLysValA 984

3035 ARCTGATTATCTTGATTTCTTGAGAGCCCTTTCAAGAGTCCAGTTC 3084  
|||||  
984 spValIleIleLeuIlePheLeuGluLysProPheGlnLysSerLysPhe 1000  
|||||  
3085 CTCACGCTCCGGAAGAGCTCTGTGGAGTTCTGTCTTGTAGTGGCCAAC 3134  
|||||  
1001 LeuGlnLeuArgLysArgLeuGlySerValLeuGluTrpProth 1017  
|||||  
3135 AAACCCGCAAGCTACCCATCTCTGGCAGTGCTTAAGAGACGCCCTGG 3184  
|||||  
1017 rAsnProGlnAlaHisProTyrPheTrpGlnCysLeuLysAsnAlaLeuA 1034  
|||||  
3185 CCACAGACAATCATGTGGCTATGTCAGGTGTTCAGGAAACGGTC 3231  
|||||  
1034 laThrAspAsnHisValAlaTyrSerGlnValPheLysGluThrVal 1049  
|||||

seq\_name: /SIDS1/gcgdata/geneseq/geneeqp-emb1/AA1999.DAT:AAW86354

seq\_documentation\_block:

ID AAW86354 standard; Protein: 1045 AA.

XX AC AAW86354;

XX 15-MAR-1999 (first entry)

XX Human DNAX toll-like receptor DTLR6.

XX DNAX toll-like receptor; DTLR: Drosophila toll receptor; IL-1 receptor;  
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;  
KW modulate inflammatory function; morphological effect;  
KW immunological disorder.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT /label= signal

FT Protein 23..1045

FT /label= DTLR6

XX WO980547-A2.

XX 12-NOV-1998.

XX 07-MAY-1998; 98WO-US08979.

XX 05-MAR-1998; 98US-0076947.

XX 07-MAY-1997; 97US-0044293.

XX 22-JAN-1998; 98US-0072212.

XX (SCHE ) SCHERING CORP.

XX Bazan JF, Hardiman GT, Kastelein RA, Rock FL;

XX WPI; 1999-059670/05.

XX N-PSDB; AAW80668.

XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter

XX phosphate metabolism, modulate inflammatory function or innate

XX immunity responses

XX Claim 5; Page 125-128; 171pp; English.

XX The present invention specifically describes human DNAX toll-like  
CC receptors 2 to 10 (DTLR2-10). The present sequence is human DTLR6  
CC given in the present invention. Also described are: (1) a fusion protein  
CC comprising a DTLR protein or peptide; (2) a binding compound, preferably  
CC an antibody or antibody fragment which specifically binds to a DTLR  
CC protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
CC peptide; (4) an expression vector comprising the nucleic acid of (3);  
CC and (5) a host cell comprising the vector of (4). The host cell of (5)  
CC can be used to produce the DTLR proteins. The DTLR proteins can be used  
CC to alter phosphate metabolism, to modulate inflammatory function, innate

CC immunity responses or morphological effects. The DTLR proteins can be  
CC used in the treatment of conditions exhibiting abnormal expression of  
CC the receptors of their ligands. These abnormalities are typically  
CC manifested by immunological disorders.

XX Sequence 1045 AA;

alignment\_scores:

Quality: 1038.00 Length: 1038

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAW86354 ..

Align seg 1/1 to: AAW86354 from: 1 to: 1045

118 ATTCTTATCTCTTTTAAACATAATCCTTAATTTCCAAACTCCTTTGGGCTAG 167

|||||

8 lleLeuIleLeuPheAsnIleIleLeuIleSerLysLeuLeuGlyAlaAr 24

168 ATGGTTCCTAAACTCTGCCCTGTGATGTCACTCTGGATGTTCACAAGA 217

|||||

24 gTrpPheProLysThrLeuProCysAspValThrLeuAspValProLysA 41

218 ACCATGTGATCGTGCACACGACAGCAAGCATTTGCACAGAAATTCCTGA 267

|||||

41 snHisValIleValAspCysThrAspLysHisLeuThrGluIleProGly 57

268 GGTATTCCAGCAACACCCAGCAACCTCACCCCTCACCATTAACACATACC 317

|||||

58 GlyIleProThrAsnThrThrAsnLeuThrLeuThrIleAsnHisIlePr 74

318 AGACATCTCCCCAGCGTCCTTTTCACAGACTGGACCATCTGGTAGATCG 367

|||||

74 oAspIleSerProAlaSerPheHisArgLeuAspHisLeuValGluIleA 91

368 ATTTTCAGATGCAACTGTGTACTATTCCACTGGGGTCAAAAACAACATG 417

|||||

91 spPheArgCysAsnCysValProIleProLeuGlySerLysAsnAsnMet 107

418 TGCATCAAGAGGCTGCAGATTAAACCCAGAGCTTTTGTAGTGGACTCCTTA 467

|||||

108 CysIleLysArgLeuGlnIleLysProArgSerPheSerGlyLeuThrTy 124

468 TTTAAATCCCTTTTACCTGGATGGAAACAGCTACTAGATACCGCAGG 517

|||||

124 rLeuLysSerLeuTyLeuAspGlyAsnGlnLeuLeuGluIleProGlnG 141

518 GCCTCCCGCTAGCTTACAGCTTCTCAGCCTTGAGGCCCAACAACATCTTT 567

|||||

141 lYleuProProSerLeuGlnLeuLeuSerLeuGluAlaAsnAsnIlePhe 157

568 TCCATCAGAAAAAGAGATCTTAACAGAACTGGCCAACTAGAAATACTCTA 617

|||||

158 SerIleArgLysGluAsnLeuThrGluLeuAlaAsnIleGluIleLeuTy 174

618 CTGGGCCAAACTGTTATTATTCGAATCCTTGTATTATGTTTCATATCAA 667

|||||

174 rLeuGlyGlnAsnCysTyTyArgAsnProCysTyValSerTySerI 191

668 TAGAGAAAGATGCCCTTCTTAACACTTGACAAAGTTAAAGTGCCTCCCTG 717

|||||

191 leGluLysAspAlaPheLeuAsnLeuThrLysLeuLysValLeuSerLeu 207

718 AAAGATAACAATGTCCACAGCCGCTCCTACTGTGTTTGGCCATCTACTTAAAC 767

|||||

208 LysaspAsnAsnValThrAlaValProThrValLeuProSerThrLeuTh 224

768 AGAACTATATCTCTACACACACATGATGCAAAAATCCAAAGAAGATGATT 817

|||||

224 rGluLeuTyLeuTyAsnAsnMetIleAlaLysIleGlnGluAspAspp 241

818 TTAATAACCTCAACAAATTACAAATCTTGACCTAAGTGGAAATGGCCCT 867  
|||||  
541 rgTyrLeuAspPheSerAsnAsnArgLeuAspLeuLeuHisSerThrAla 557  
|||||  
1768 TTTCAGAGCTTCACAACTCGAAGTTCTCGATATACACGACGTAATAGCCA 1817  
|||||  
558 PheGluGluLeuHisLysLeuGluValLeuAspIleSerSerAsnSerHI 574  
|||||  
1818 TTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTTACCAGAACC 1867  
|||||  
574 sTyrPheGlnSerGluGlyIleThrHisMetLeuAsnPhethrLysAsnL 591  
|||||  
1868 TAAAGTTCTGCAGAACTGATGATGAACAGCAATGACATCTCTTCCTCC 1917  
|||||  
591 euLysValLeuGlnLysLeuMetMetAsnAspAsnAspIleSerSerSer 607  
|||||  
1918 ACCAGCAGGACCATGGAGAGTGCAGTCTCTTAGAACTCTGGAATTCAGAGG 1967  
|||||  
608 ThrSerArgThrMetGluSerGluSerLeuArgThrLeuGluPheArgGL 624  
|||||  
1968 AAATCACCTTAGATGTTTATGGAGAGAAGGTGATAACAGATACTTACAAT 2017  
|||||  
624 yAsnHisLeuAspValLeuTprArgGluGlyAspAsnArgTyrLeuGlnL 641  
|||||  
2018 TATTCAGAATCTGCTAAATTTAGAGGAATTAGACATCTCTAAATAATTC 2067  
|||||  
641 euPheLysAsnLeuLeuLysLeuGluGluLeuAspIleSerLysAsnSer 657  
|||||  
2068 CTAAGTTCTTGGCCTTCGGAGTTTTCATGGTATGCTCCCTCCAAATCTAAA 2117  
|||||  
658 LeuSerPheLeuProSerGlyValPheAspGlyMetProAsnLeuLys 674  
|||||  
2118 GAATCTCTCTTTGGCCAAAATGGGCTCAAATCTTTTCAGTTTGGGAAGAAC 2167  
|||||  
674 sasLeuSerLeuAlaLysAsnGlyLeuLysSerPheSerTrpLysLysL 691  
|||||  
2168 TCCAGTGTCTAAAGAACCTGGAAACTTTGGACCTCAGCCACAAACCACTG 2217  
|||||  
691 euGlnCysLeuLysAsnLeuGluThrLeuAspLeuSerHisAsnGlnLeu 707  
|||||  
2218 ACCACTGTCCCTGAGAGATTATCCAACTGTTCCAGAAAGCTCAAGAATCT 2267  
|||||  
708 ThrThrValProGluArgLeuSerAsnCysSerArgSerLeuLysAsnLe 724  
|||||  
2268 GATTCCTTAAGAATAATCAAACTCAGGAGTCTGACGAAGTATTTCTACAAG 2317  
|||||  
724 uIleLeuLysAsnAsnGlnIleArgSerLeuThrLysTyrPheLeuGlnA 741  
|||||  
2318 ATGCTTCCAGTTGCGATATCTGGATCTCAGCTCAATATAAATCCAGATG 2367  
|||||  
741 spAlaPheGlnLeuArgTyrLeuAspLeuSerSerAsnLysIleGlnMet 757  
|||||  
2368 ATCCAAAAGACGAGCTTCCCAGAAAATGCTCAACAATCTGAAGATGTT 2417  
|||||  
758 IleGlnLysThrSerPheProGluAsnValLeuAsnAsnLeuLysMetLe 774  
|||||  
2418 GCTTTTGCATCATATCGTTTCTGTCACCTGTGATGCTGTGTGGTTTG 2467  
|||||  
774 uLeuLeuHisLysAsnArgPheLeuCysThrCysAspAlaValTrpPheV 791  
|||||  
2468 TCTGTGGGTTAAACATACGAGGAGTGACTATTCTCTTACCTGGCCACAGAT 2517  
|||||  
791 alTrpTrpValAsnHisThrGluValThrIleProTyrLeuAlaThrAsp 807  
|||||  
2518 GTGACTTGTGGGCCAGGAGCACACAAGGCCAAAGTGTGATCTCCCT 2567  
|||||  
808 ValThrCysValGlyProGlyAlaHisLysGlyGlnSerValIleSerLe 824  
|||||  
2568 GGATCTGTACACCTGTGAGTTAGATCTGACTAACTGATCTCTGTCTCTCAC 2617  
|||||  
824 uAspLeuTyrThrCysGluLeuAspLeuThrAsnLeuIleLeuPheSerL 841  
|||||  
2618 TTTCCATATCTGTATCTCTCTTCTCATGTGATGATGACAGCAAGTCAC 2667  
|||||

1018 AACATCAACAACCTCCAGGAAGTGGATCTGCTCCCAAACTCTTGGCCAA 1067  
|||||  
308 AsnIleAsnLysLeuGlnGluLeuAspLeuSerGlnAsnPheLeuAlaLy 324  
|||||  
1068 AGAAATTGGGATGCTAAATTTCTGCAATTTCTCCCGAGCTCATCCAAT 1117  
|||||  
324 sGluIleGlyAspAlaLysPheLeuHisPheLeuProSerLeuIleGlnL 341  
|||||  
1118 TGGATCTGCTCTTCAATTTTGAACCTTCAGGCTATCGTGCATCTATGAAT 1167  
|||||  
341 euAspLeuSerPheAsnPheGluLeuGlnValTyrArgAlaSerMetAsn 357  
|||||  
1168 CTATCACAAGCATTTCTTCTACTGAAAAGCTGAAAATCTCGCGGATCAG 1217  
|||||  
358 LeuSerGlnAlaPheSerSerLeuLysSerLeuLysIleLeuArgIleAr 374  
|||||  
1218 AGGATATGCTTTTAAAGAGTTGAAAAGCTTTAACTCTCGCCATTTACATA 1267  
|||||  
374 gGlyTyrValPheLysGluLeuLysSerPheAsnLeuSerProLeuHisA 391  
|||||  
1268 ATCTTCAAAATCTTGAAGTCTTGATCTTGGCACTTAACCTTATRAAAAT 1317  
|||||  
391 snLeuGlnAsnLeuGluValLeuAspLeuGlyThrAsnPheIleLysIle 407  
|||||  
1318 CCTAACCTCAGCATGTTTAAACAATTTAAAAGACTGAAAGTCATAGATCT 1367  
|||||  
408 AlaAsnLeuSerMetPheLysGlnPheLysArgLeuLysValIleAspLe 424  
|||||  
1368 TTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCT 1417  
|||||  
424 uSerValAsnLysIleSerProSerGlyAspSerSerGluValGlyPheC 441  
|||||  
1418 GCTCAAAATGCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCCTGGAA 1467  
|||||  
441 ysSerAsnAlaArgThrSerValGluSerTyrGluProGlnValLeuGlu 457  
|||||  
1468 CAATTTACATTATTTTCAGATATGATGAAGTATGCAAGGAGTTCAGATTCAA 1517  
|||||  
458 GlnLeuHisLysTyrPheArgTyrAspLysTyrAlaArgSerCysArgPheLy 474  
|||||  
1518 AAACAAGAGGCTCTTTTCATGCTCTGTTTAATGAAAGCTGTACAAGTATG 1567  
|||||  
474 sasLysGluAlaSerPheMetSerValAsnGluSerCysTyrLysTyrG 491  
|||||  
1568 GCGAGACCTTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCCTCT 1617  
|||||  
491 yGlnThrLeuAspLeuSerLysAsnSerIlePhePheValLysSerSer 507  
|||||  
1618 GATTTTTCAGCATCTTCTTCTTCTCAAAATGCTGATCTGTGAGGAAATCT 1667  
|||||  
508 AspPheGlnHisLeuSerPheLeuLysCysLeuAsnLeuSerGlyAsnLe 524  
|||||  
1668 CATTAGCCAAACTCTTAATGGCAGTGAATTTCCAACCTTTTAGCAGAGCTGA 1717  
|||||  
524 uIleSerGlnThrLeuAsnGlySerGluPheGlnProLeuAlaGluLeuA 541  
|||||

```
|||||
841 euSerIleSerValSerLeuPheLeuMetValMetMetThAlaSerHis 857
2668 CTTCTATTTCTGGATGTGGTATATTTACCATTTCTGTAAAGCCCAAGAT 2717
|||||
858 LeuTyPheTrpAspValTrpPyrIleTyHISPhECysLysAlaLysIle 874
2718 AAAGGGCTATCAGCTCTAAATATCACAGACTGTTGCTATGATGCTTTTA 2767
|||||
874 elySgLyTyGlnAglLeuIleSerProAspCysTyTrpAspAlaPheI 891
2768 TTGTGTATGACACTAAAGACCCAGCTGTGACCGAGTGGGTTTGGCTGAG 2817
|||||
891 leValTyAspThrLysAspProAlaValThrGluTrpValLeuAlaGlu 907
2818 CNGGTGGCCAACTGGAAGACCCAGACAGAGAAACATTTTAAATTATGTCT 2867
|||||
908 LeuValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCysLe 924
2868 CGAGGAAAGGGACTGTTTACCAGGCGAGCCAGTCTCGAAAACCTTTGCC 2917
|||||
924 uGluGluArgAspTrpLeuProGlyGlnProValLeuGluAsnLeuSerG 941
2918 AGAGCATACAGCTTAGCAAAAGACAGTGTGTTGTGATGACAGACAAGTAT 2967
|||||
941 InSerIleGlnLeuSerLysLysThrValPheValMetThrAspLysTy 957
2968 GCAAGAGCTGAAATTTTAAAGATACATTTTACTTGTCCCATCAGAGCT 3017
|||||
958 AlaLysThrGluAsnPhelyIleAlaPheTyLeuSerHisGlnArgLe 974
3018 CATGGATGAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAAGCCT 3067
|||||
974 uMetAspGluLysValAspValIleIleLeuIlePheLeuGluLysProp 991
3068 TTCAGAAGTCCAAAGTCTCTCCAGCTCCGGAAAGCCTCTGTGGGAGTTCT 3117
|||||
991 heGlnLysSerLysPheLeuGlnLeuArgLysArgLeuCysGlySerSer 1007
3118 GTCCCTTGAGTGCCCAACAAACCCGCAAGCTCACCCATCTCTTGGCAGTG 3167
|||||
1008 ValLeuGluTrpProThrAsnProGlnAlaHisProTyPheTrpGlnCy 1024
3168 TCTAAAGACGCCCTGGCCACACACAATCATGTGGCCCTATAGTCAGGTGT 3217
|||||
1024 sLeuLysAsnAlaLeuAlaThrAspAsnHisValAlaTySerGlnValP 1041
3218 TCAAGGAAACGGTC 3231
|||||
1041 heLysGluThrVal 1045
```

seq\_name: /SID1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAW58342

```
seq_documentation_block:
ID- AAW58342 standard; Protein; 432 AA.
XX
AC AAW58342;
XX
XX 05-NOV-2001 (first entry)
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30447.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapIens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
```

```
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 30447; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 432 AA;
```

```
alignment_scores:
Quality: 432.00 Length: 432
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-202-054-2 x AAW58342 ..
Align seg 1/1 to: AAW58342 from: 1 to: 432
```

```
88 GTGTTTCCAATGTGGACACTGAAGACAAATCTTATCTCTTTTAAACAT 137
|||||
1 ValPheProMetTrpThrLeuLysArgGlnIleLeuIleLeuPheAsnIle 17
138 AATCCTAATTTCCAAACTCCTTGGGGCTAGATGGTTTCTTAAACTCTGC 187
|||||
17 eileLeuIleSerLysLeuLeuGlyAlaArgTrpPheProLysThrLeuP 34
188 CCTGTGATGTCACTCTGATGTTCCAAAGAACCATGTGATCTGGACTGC 237
|||||
34 rOCysAspValThrLeuAspValProLysAsnHisValIleValAspCys 50
238 ACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTTCCACAGAACACCAC 287
|||||
51 ThrAspLysHisLeuThrGluIleProGlyGlyIleProThrAsnThrTh 67
288 GAACCTCACCTCACCATTACCACATACCAGACATCTCCCCAGGCTCCT 337
|||||
67 rAsnLeuThrLeuThrIleAsnHisIleProAspIleSerProAlaSerP 84
338 TTCACAGACTCGACCATCTGGTAGAGATCGATTTTCAGATGCAACTGTGTA 387
|||||
84 heHisArgLeuAspHisLeuValGluIleAspPheArgCysAsnCysVal 100
388 CCTATTCCACTGGGGTCAAAAACAAACATGTGCATCAAGAGGCTGCAGAT 437
|||||
101 ProIleProLeuGlySerLysAsnAsnMetCysIleLysArgLeuGlnIle 117
438 TAAACCCAGAACGCTTACTGGACTCAGCTTATTTAAATCCCTTTACCTGG 487
|||||
117 elysProArgSerPheSerGlyLeuThrTyLeuLysSerLeuTyLeuA 134
```

488	ATGGAACACGCTACTAGATACCCAGCGCTCCCGCTAGCTTACAG	537	seq_documentation_block:
134	spGlyAsnGlnLeuLeuGluIleProGlnGlyLeuProSerLeuGln	150	ID AAM43566 standard; Protein; 235 AA.
538	CTTCTCAGCCTTGAGGCCAACACATCTTTTCCATCAGAAAGAGAACTCT	587	XX AAM43566;
151	LeuLeuSerLeuGluAlaAsnAsnIlePheSerIleArgLysGluAsnLe	167	XX 22-OCT-2001 (first entry)
588	AACAGAACTGGCCACATAGAATACTCTACCTGGGCCCAAAACTGTTATT	637	XX Human polypeptide SEQ ID NO 244.
167	uthrGluLeuAlaAsnIleGluLeuTyrLeuGlyGlnAsnCystYrT	184	XX Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
638	ATCGAAATCCTTGTATTGTTTCATATTCATAGAGAAAGATCGCTTCCPA	687	KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
184	YrArgAsnProCysTyrValSerTyrSerIleGluLysAspAlaPheLeu	200	KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
688	AACTTCACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGC	737	KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
201	AsnLeuThrLysLeuLysValLeuSerLysAspAsnValThrAl	217	KW antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;
738	CGTCCCTACTGTTTGGCCATCTACTTTAACAGAACTATATCTCTACAACA	787	KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;
217	aValProThrValLeuProSerThrLeuThrGluLeuTyrLeuTyrAsna	234	KW cardiovascular disorder; neurological disease; infection; human.
788	ACATGATTGCAAAAATCCAAGAAGATGATTTTAATAACCTCAACCAATTA	837	XX Homo sapiens.
234	snMetIleAlaLysIleGlnGluAspAspPheAsnAsnLeuAsnGlnLeu	250	XX WO200155308-A2.
838	CAAAATCTTGACCTAGTGAATTCCTCCCTCGTGTGTTATTAATGCCCATTT	887	XX 02-AUG-2001.
251	GlnIleLeuAspLeuSerGlyAsnCysProArgCysTyrAsnAlaProPh	267	XX 17-JAN-2001; 2001WO-US01309.
888	TCTCTGCGCCGTGAAAAATTAATTTCTCCCTACAGATCCCTGTAATG	937	PR 31-JAN-2000; 2000US-0179065.
267	eProCysAlaProCysLysAsnAsnSerProLeuGlnIleProValAsna	284	PR 04-FEB-2000; 2000US-0180628.
938	CTTTTGATGGGTGACAGAATTAAGTTTTTACGTCTACACAGTAAGTCT	987	PR 24-FEB-2000; 2000US-0184664.
284	laPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsnSer	300	PR 02-MAR-2000; 2000US-0186350.
988	CTTCACAGTGTGCCCCAGAGTGGTTTAAGACATCAACAACTCCAGGA	1037	PR 16-MAR-2000; 2000US-0189874.
301	LeuGlnHisValProArgIrrPheLysAsnIleAsnLysLeuGlnG	317	PR 17-MAR-2000; 2000US-0190076.
1038	ACTGATCTGTCCCAAACTTCTTGCCCAAGAAATTTGGGGATGCTAAAT	1087	PR 18-APR-2000; 2000US-0198123.
317	uLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaLysP	334	PR 19-MAY-2000; 2000US-0205515.
1088	TTCTGATTTTCTCCCGCCTCATCCAAATGGATCTGCTTTCAATTTT	1137	PR 07-JUN-2000; 2000US-0209467.
334	heLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsnPhe	350	PR 28-JUN-2000; 2000US-0214886.
1138	GAACTTCAGTCTATCGTCATCTATGAATCTATCACAAGCATTTTCTTC	1187	PR 30-JUN-2000; 2000US-0215135.
351	GluLeuGlnValTyrArgAlaSerMetAsnLeuSerGlnAlaPheSerSe	367	PR 07-JUL-2000; 2000US-0216647.
1188	ACTGAAAGCCTGAAATTTCTCGGATCAGAGATATGTCTTTAAAGAGT	1237	PR 11-JUL-2000; 2000US-0217487.
367	rLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysGluL	384	PR 11-JUL-2000; 2000US-0217496.
1238	TGAAAGCTTTAACCTCTCGCCATTCATAATCTTCAAAATCTTGAAGTT	1287	PR 14-JUL-2000; 2000US-0218290.
384	eulysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGluVal	400	PR 26-JUL-2000; 2000US-0220963.
1288	CTTGATCTGGCACTAACTTTATAAAATTCGTAACCTCAGCATGTTTAA	1337	PR 14-AUG-2000; 2000US-0224518.
401	LeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPheLy	417	PR 14-AUG-2000; 2000US-0225266.
1338	ACAAATTTAAAGACTGAAAGTCATAGATCTTTCAAGTGAATAAATA	1383	PR 14-AUG-2000; 2000US-0225267.
417	sGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIle	432	PR 14-AUG-2000; 2000US-0225267.
seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAM43566			PR 14-AUG-2000; 2000US-0225268.

PR 08-SEP-2000; 2000US-02314113.  
PR 08-SEP-2000; 2000US-02314114.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-488781/53.  
DR N-PSDB; AAI63872.  
XX  
PT New isolated nucleic acids and polypeptides, useful for diagnosing,  
PT treating and/or preventing human diseases and disorders -  
XX  
PS Claim 11; SEQ ID NO 244; 664pp + Sequence Listing; English.  
XX  
CC The invention relates to human polynucleotides (AAI63803-AAI64012) and  
CC the encoded proteins (AAM434497-AAM43660) useful for preventing, treating  
CC or ameliorating medical conditions e.g. by protein or gene therapy. The  
CC genes were isolated from a range of human tissues disclosed in the  
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists  
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,  
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,  
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or  
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,  
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 235 AA;  
  
alignment\_scores:  
Quality: 233.00 Length: 233  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x AAM43566 ..  
  
Align seg 1/1 to: AAM43566 from: 1 to: 235  
  
2470 TGGTGGTTAAACATACGGAGGTGACTATTCTTACCTGGCCACAGATGT 2519  
|||||  
3 TrpTrpValAsnHisThrGluValThrIleProTyrLeuAlaThrAspVa 19  
  
2520 GACTTGTGTGGGGCCAGGAGCACACAAAGGCCAAAGTGTGATCTCCCTGG 2569  
|||||  
19 lThrCysValGlyProGlyAlaHisLysGlyGlnSerValIleSerLeuA 36  
  
2570 ATCTGTACACCTGTGAGTTAGATCTGACTTAACCTGATTCTGTTCACATT 2619  
|||||  
36 splenTyThrCysGluLeuAspLeuThrAsnLeuIleLeuPheSerLeu 52

2620 TCATATCTGTATCTCTCTTTCTCATGTGATGACAGCAAGTCACT 2669  
|||||  
53 SerIleSerValSerLeuPheLeuMetValMetMetThrAlaSerHisLe 69  
2670 CTATTTCTGGGATGCTGTATATTACCATTTCTGAAGGCCAAGATAA 2719  
|||||  
69 utyPheTrpAspValTrpTyrHisPheCysLysAlaLysIleL 86  
2720 AGGGTATCAGCGTCTAATATATCACCAGACTGTTGCTATGATGCTTTATT 2769  
|||||  
86 ysgLtyrGlnargLeuIleSerProAspCysTyrAspAlaPheIle 102  
2770 GGTATGACACATAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCT 2819  
|||||  
103 ValTyrAspThrLysAspProAlaValThrGluTrpValLeuAlaGluLe 119  
2820 GTGGCCAACTGGAGAGCCCAAGACAGAAACATTTAATTTATGCTCTCG 2869  
|||||  
119 uValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCysLeuG 136  
2870 AGGAAAGGACTGGTTTACCAGGCGCAGCTTCTGTGAAAACCTTTCCCGAG 2919  
|||||  
136 luGluArgAspTrpLeuProGlyGlnProValLeuGluAsnLeuSerGln 152  
2920 AGCATACAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGC 2969  
|||||  
153 SerIleGlnLeuSerLysLysThrValPheValMetThrAspLysTyrAl 169  
2970 AAAGACTGAAATTTTAAGATAGCATTTTACTTGTCCTCCCATCAGAGCTCA 3019  
|||||  
169 aLysThrGluAsnPheLysIleAlaPheTyrLeuSerHisGlnArgLeuM 186  
3020 TGCATGAAAAAGTTGATGTGATTATCTTCATATTTCTTGAGAGGCCCTTT 3069  
|||||  
186 etAspGluLysValaspValIleIleLeuIlePheLeuGluLysProPhe 202  
3070 CAGAACTCAAGTTCCTCCAGCTCCGGAAGGCTCTGTGGGAGTTCTGT 3119  
|||||  
203 GlnLysSerLysPheLeuGlnLeuArgLysArgLeuCysGlySerSerVa 219  
3120 CCTGAGTGGCCAAACCCGCAAGCTCACCATACTTCTGGCAGCTGT 3168  
|||||  
219 lLeuGluTrpProThrAsnProGlnAlaHisProTyrPheTrpGlnCys 235

seq\_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAU17407

seq\_documentation\_block:

ID AAU17407 standard; Protein; 161 AA.

XX AC AAU17407;

XX DT 07-NOV-2001 (first entry)

XX DE Novel signal transduction pathway protein, Seq ID 972.

XX KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS;  
KW acquired immune deficiency syndrome.

XX OS Homo sapiens.

XX PN WO200154733-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01312.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241285.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-465460/50.  
DR N-PSDB; AAS27324.  
XX  
PT Novel polypeptides useful for diagnosing, treating, preventing and/or

PT prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders -  
PS Claim 1; SEQ ID No 972; 880pp; English.  
XX  
CC The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative  
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
CC Addison's disease), reproductive system disorders, gastrointestinal  
CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
CC as stimulators of B-cell responsiveness to pathogens, activators of  
CC T-cells, to induce higher affinity antibodies, and as a means to induce  
CC tumour proliferation in pathologies e.g. acquired immune deficiency  
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction  
CC pathway protein, amino acid sequences of the invention.  
XX

alignment\_scores:

Quality: 161.00 Length: 161  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAU17407 ..

Align seg 1/1 to: AAU17407 from: 1 to: 161

2470 TGGTGGTTAACCATACGAGGTGACTATTCTTACCTGGCCACAGATGT 2519  
1 TPTTPVAIASNHSTHRRGluValThrIleProTyrLeuAlaThrAspVa 17  
2520 GACTTGTGTGGGCCAGCAGCACACAAAGGCCAAAGTGTGATCTCCCTGG 2569  
17 lThrCysValGlyProGlyAlaHisLysGlyGlnSerValIleSerLeuA 34  
2570 ATCTGTACACCTGTGAGTTAGATCTGACTTAACCTGATTCTGTTCACATT 2619  
34 sPLeuTyrThrCysGluLeuAspLeuThrAsnLeuIleLeuPheSerLeu 50  
2620 TCCATATCTGATCTCTCTTCTCATGGTGATCATCAGCAGCAAGTCCCT 2669  
51 SerIleSerValSerLeuPheLeuMetValMetMetThrAlaSerHisLe 67  
2670 CTATTTCTGGGATGTGTGTATATTTTACCATTTCTTAAGGCCAAGATAA 2719  
67 uTyrPheTrpAspValTrpTyrIleTyrHisPheCysLysAlaLysIleL 84  
2720 AGGGGTATCAGGCTCTAATATCACCAGACTGTTGCTATGATGCTTTATT 2769  
84 ysGlyTyrGlnArgLeuIleSerProAspCysCystyrAspAlaPheIle 100  
2770 GTGTATGACACTAAAGACCCAGCTGTGACCGAGTGGTTTTTGGCTGAGCT 2819  
101 ValTyrAspThrLysAspProAlaValThrGluTrpValLeuAlaGluLe 117  
2820 GGTGCCCAAACTGAGAGCCCAAGAGAGAAACATTTTAAATTTATGTCTCG 2869  
117 uValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCysLeuG 134  
2870 AGAAAGGACTGTTTACCAGGCGCAGTCTCTTGAAACCTTTCCCGAC 2919  
|||||



134 luGluArgAspTrpLeuProGlyGlnProValLeuGluAsnLeuSerGln 150  
2920 AGCATACAGCTTACGAAAGACAGTGTGTG 2952  
|||||  
151 SerIleGlnLeuSerIlysthrValPheVal 161

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AA43641

seq\_documentation\_block:

ID AA43641 standard; Protein; 161 AA.

AC AA43641;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 319.

XX

KW Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;  
cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;  
neuroprotective; antiallergic; hepatotropic; antidiabetic;  
antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;  
antiparasitic; cardiant; gene therapy; cancer; immune disorder;  
cardiovascular disorder; neurological disease; infection; human.

XX

OS Homo sapiens.

XX

PN WO200155308-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01309.

XX

31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2001US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488781/53.

DR N-PSDB; AAI63947.

XX New isolated nucleic acids and polypeptides, useful for diagnosing,  
PT treating and/or preventing human diseases and disorders -

XX Claim 11; SEQ ID NO 319; 664pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI63803-AAI64012) and  
CC the encoded proteins (AAM43497-AAM43660) useful for preventing, treating  
CC or ameliorating medical conditions e.g. by protein or gene therapy. The  
CC genes were isolated from a range of human tissues disclosed in the  
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists  
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,  
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,  
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or  
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,  
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 161 AA;

alignment\_scores:  
Quality: 161.00 Length: 161  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAM43641 ..

Align seg 1/1 to: AAM43641 from: 1 to: 161

2470 TGGTGGGTTAAACCATCGGAGGTGACTATTCCTTACCTGCCACAGATGT 2519  
1 TrrprvalasnHisThrGluValThrIleProTyrLeuAlaThrAspVa 17

2520 GACTTGTGGGGCCAGGAGCACACAAAGGCCAAAGTGTATCTCCCTGG 2569

117 lThrCysValGlyProGlyAlaHisLysGlyGlnSerValIleSerLeuA 34  
2570 ATCTGTACACCTGTGAGTTAGATCTGACTAACTGATTCTGTCTCCTCATT 2619  
34 spluTyrThrCysGluLeuAspLeuThrAsnLeuIleLeuPheSerLeu 50  
2620 TCCATATCTGTATCTCTCTTCTCATGGTGATGACACACCAAGTCCACCT 2669  
51 SerIleSerValSerLeuPheLeuMetValMetMetThrAlaSerHisLe 67  
2670 CTATTCTGGGATGTGGTATATTTACCATTTCTGTAAAGGCCAAGATAA 2719  
67 uTyrPheTrpAspValTrpTyrIleTyrHisPheCysLysAlaLysIleL 84  
2720 AGGGGTATCAGCGCTAAATATACACAGACTGTTGCTATGATGCTTTTATT 2769  
84 ysglyTyrGlnArgLeuIleSerProAspCysCysTyrAspAlaPheIle 100  
2770 GTGTATGACACTAAAGACCCAGCTGTGACGAGTGGGTTTGGCTGAGCT 2819  
101 ValTyrAspThrLysAspProAlaValThrGluTrpValLeuAlaGluLe 117  
2820 GTGGCCAAACTCGAAGACCCAAAGAGAGAAACATTTTAAATTTATGCTCG 2869  
117 uValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCysLeuG 134  
2870 AGGAAAGGACTGTTTACACAGGCAGCCAGTCTCTGGAAAACCTTTCCAG 2919  
134 luGluArgAspTrpLeuProGlyGlnProValLeuGluAsnLeuSerGln 150  
2920 AGCATACAGCTTACCAAAAGACAGTGTGTGTG 2952  
151 SerIleGlnLeuSerLysThrValPheVal 161

seq\_name: /sids1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAM90355

seq\_documentation\_block:

ID AAM90355 standard; Protein; 121 AA.

XX AC AAM90355;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen SEQ ID NO:17948.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX KW cytostatic; gene therapy; vaccine; metastasis.

XX OS Homo sapiens.

XX PN WO2000157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01354.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 30-JUN-2000; 2000US-0214886.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 11-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 14-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218290.



CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 121 AA;

alignment\_scores:  
Quality: 64.00 Length: 78  
Ratio: 0.831 Gaps: 1  
Percent Similarity: 98.718 Percent Identity: 98.718

alignment\_block:  
US-09-202-054-2 x AAM90355 ..  
  
Align seg 1/1 to: AAM90355 from: 1 to: 121  
  
456 TGGACTCACTATTATAAATCCCTTTACCTGGATGGAACACGACTACTAG 505  
|||||  
1 TrpThrHisLeuPheLysIleProLeuProGlyTyrLysProAlaThrAr 17  
  
506 AGATACCGCAGGCGCTCCCGCTAGCTTACAGCTTCTCAGCCTTGAGGCC 555  
|||||  
17 GAspThrAla.GlyLeuProSerLeuGlnLeuLeuSerLeuGluAla 33  
  
556 AACACATCTTTCCATCATCAGAAAGAGATCTAACAGAACTGGCCCAACAT 605  
|||||  
34 AsnAsnIlePheSerIleArgLysGluAsnLeuThrGluLeuAlaAsnIi 50  
  
606 AGAATACTCTACCTGGGCAAACTGTTATTATCGAAATCTCTGTTATG 655  
|||||  
50 eGluLeuLeuTyrLeuGlyGlnAsnCysTyrTyrArgAsnProCysTyrV 67  
  
656 TTTCATATTCAATAGAGAAAGATGCTTCCTA 687  
|||||  
67 alserTyrSerIleGluLysAspAlaPheLeu 77

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAW86363

seq\_documentation\_block:  
ID -AAW86363 standard; Protein: 394 AA.

XX AAW86363;  
AC  
XX  
XX  
DT 15-MAR-1999 (first entry)  
XX  
XX Mouse DNAX toll-like receptor DTLR6.  
XX  
XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;  
XX interleukin 1 receptor; phosphate metabolism; innate immunity response;  
XX modulate inflammatory function; morphological effect;  
XX immunological disorder.  
XX  
XX Mus sp.

XX WO9805047-A2.  
XX  
XX 12-NOV-1998.  
XX  
XX  
XX 07-MAY-1998; 98WO-US08979.  
XX  
XX 05-MAR-1998; 98US-0076947.  
XX 07-MAY-1997; 97US-0044293.  
XX 22-JAN-1998; 98US-0072212.  
XX  
XX (SCHE ) SCHERING CORP.  
XX  
XX Bazan JF, Hardiman GT, Kastelein RA, Rock FL;  
XX  
XX WPI; 1999-059670/05.  
XX N-PSDB; AAV80677.  
DR

XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter  
PT phosphate metabolism, modulate inflammatory function or innate  
PT immunity responses  
XX  
XX Example; Page 154-155; 171pp; English.  
XX  
XX The present invention specifically describes human DNAX toll-like  
CC receptors 2 to 10 (DTLR2-10). The present sequence is mouse DTLR6  
CC given in the present invention. Also described are: (1) a fusion  
CC protein comprising a DTLR protein or peptide; (2) a binding compound,  
CC preferably an antibody or antibody fragment which specifically binds to  
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
CC peptide; (4) an expression vector comprising the nucleic acid of (3);  
CC and (5) a host cell comprising the vector of (4). The host cell of (5)  
CC can be used to produce the DTLR proteins. The DTLR proteins can be used  
CC to alter phosphate metabolism, to modulate inflammatory function, innate  
CC immunity responses or morphological effects. The DTLR proteins can be  
CC used in the treatment of conditions exhibiting abnormal expression of  
CC the receptors of their ligands. These abnormalities are typically  
CC manifested by immunological disorders.  
XX  
SQ Sequence 394 AA;

alignment\_scores:  
Quality: 48.00 Length: 48  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAW86363 ..  
  
Align seg 1/1 to: AAW86363 from: 1 to: 394

2815 GAGCTGGTGGCCAAACTGGAGACCCAGACAGAAACATTTTAAATTATG 2864  
|||||  
256 GluLeuValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCy 272  
  
2865 TCTCAGGAAAGGACTGGTTACACGGGAGCCAGCTTCGGAAACCTTT 2914  
|||||  
272 sLeuGluGluArgAspTyrLeuProGlyGlnProValLeuGluAsnLeuS 289  
  
2915 CCAGAGCATACAGCTTAGCAAAAGACAGCTGTTGTGATGACA 2958  
|||||  
289 eGlnSerIleGlnLeuSerLysLysThrValPheValMetThr 303

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAU18552

seq\_documentation\_block:  
ID -AAU18552 standard; Protein: 101 AA.

XX AAU18552;  
AC  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX Human lung antigen polypeptide #1.

XX Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;  
XX chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
XX antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
XX cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
XX ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
XX hyperproliferative disorder; breast; liver; cardiovascular disorder;  
XX cerebrovascular disorder; nervous system disorder; bacterial infection;  
XX fungal infection; viral infection; ocular disorder; endocrine disorder;  
XX gastrointestinal disorder; renal disorder; respiratory disorder;  
XX wound healing; skin aging; organ transplantation; food preservative;  
XX tissue regeneration; anti-infertility; food additive.  
XX  
XX Homo sapiens.  
XX  
XX WO200155303-A2.  
PN

XX PD 02-AUG-2001.  
XX XX 17-JAN-2001; 2001WO-US01301.  
XX XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX XX (HUMA-) HUMAN GENOME SCI INC.  
XX XX Rosen CA, Barash SC, Ruben SM;  
PI XX

```
DR WPI; 2001-457723/49.
XX N-PSDB; AAS29839.
XX Isolated polypeptide for treating, preventing and/or prognosing
PT respiratory disorders related to the lung including lung cancers and
PT also for testing and detection e.g. diagnosis -
XX
PS Claim 11; SEQ ID NO 103; 507pp; English.
XX
XX Sequences AAS29839-AAS29930 represent the lung antigen polypeptides of
CC the invention. Lung antigen polypeptides and their associated
CC polynucleotides are useful in the diagnosis, treatment and prevention of
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,
CC cats, dogs, chickens or sheep. A pathological condition can be determined
CC by detecting the presence or absence of a mutation in a lung antigen
CC polynucleotide. The treatable disorders include autoimmune diseases such
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as premature labour, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

alignment_scores:
Quality: 28.00 Length: 28
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAU18552 ..
Align seg 1/1 to: AAU18552 from: 1 to: 101
3 CATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGA 52
|||||
5 HisLeuLysLeuIleLeuAlaProLeuMetLeuCysSerLeuGlnProAs 21
53 CCTCTACATCCATTCTTGGAGAGACTAAAAAT 86
|||||
21 pLeuTyrlleProPheTrpLysThrLysAsn 32
|||||

seq_name: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:AAW86355
seq_documentation_block:
ID AAW86355 standard; Protein; 59 AA.
XX
XX AAW86355;
XX
XX 15-MAR-1999 (first entry)
XX
XX Partial mouse DNAX toll-like receptor DTLR6.
XX
XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;
KW modulate inflammatory function; morphological effect;
KW immunological disorder.
XX
XX Mus sp.
XX
XX WO9850547-A2.
XX
XX 12-NOV-1998.
XX
XX 07-MAY-1998; 98WO-US08979.

DR WPI; 2001-457723/49.
XX N-PSDB; AAS29839.
XX Isolated polypeptide for treating, preventing and/or prognosing
PT respiratory disorders related to the lung including lung cancers and
PT also for testing and detection e.g. diagnosis -
XX
PS Claim 11; SEQ ID NO 103; 507pp; English.
XX
XX Sequences AAS29839-AAS29930 represent the lung antigen polypeptides of
CC the invention. Lung antigen polypeptides and their associated
CC polynucleotides are useful in the diagnosis, treatment and prevention of
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,
CC cats, dogs, chickens or sheep. A pathological condition can be determined
CC by detecting the presence or absence of a mutation in a lung antigen
CC polynucleotide. The treatable disorders include autoimmune diseases such
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as premature labour, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

alignment_scores:
Quality: 28.00 Length: 28
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAU18552 ..
Align seg 1/1 to: AAU18552 from: 1 to: 101
3 CATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGA 52
|||||
5 HisLeuLysLeuIleLeuAlaProLeuMetLeuCysSerLeuGlnProAs 21
53 CCTCTACATCCATTCTTGGAGAGACTAAAAAT 86
|||||
21 pLeuTyrlleProPheTrpLysThrLysAsn 32
|||||

seq_name: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:AAW86355
seq_documentation_block:
ID AAW86355 standard; Protein; 59 AA.
XX
XX AAW86355;
XX
XX 15-MAR-1999 (first entry)
XX
XX Partial mouse DNAX toll-like receptor DTLR6.
XX
XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;
KW modulate inflammatory function; morphological effect;
KW immunological disorder.
XX
XX Mus sp.
XX
XX WO9850547-A2.
XX
XX 12-NOV-1998.
XX
XX 07-MAY-1998; 98WO-US08979.

XX 05-MAR-1998; 98US-0076947.
PR 07-MAY-1997; 97US-0044293.
PR 22-JAN-1998; 98US-0072212.
XX
XX (SCHE ) SCHERING CORP.
XX
XX Bazan JF, Hardiman GT, Kastelein RA, Rock FL;
XX
XX WPI; 1999-059670/05.
DR N-PSDB; AAW80669.
XX
XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter
PT phosphate metabolism, modulate inflammatory function or innate
PT immunity responses
XX
XX Example; Page 129; 171pp; English.
XX
XX The present invention specifically describes human DNAX toll-like
CC receptors 2 to 10 (DTLR2-10). The present sequence is partial mouse
CC DTLR6 given in the present invention. Also described are: (1) a fusion
CC protein comprising a DTLR protein or peptide; (2) a binding compound,
CC preferably an antibody or antibody fragment which specifically binds to
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or
CC peptide; (4) an expression vector comprising the nucleic acid of (3);
CC and (5) a host cell comprising the vector of (4). The host cell of (5)
CC can be used to produce the DTLR proteins. The DTLR proteins can be used
CC to alter phosphate metabolism, to modulate inflammatory function, innate
CC immunity responses or morphological effects. The DTLR proteins can be
CC used in the treatment of conditions exhibiting abnormal expression of
CC the receptors of their ligands. These abnormalities are typically
CC manifested by immunological disorders.
XX
XX Sequence 59 AA;

alignment_scores:
Quality: 16.00 Length: 16
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAW86355 ..
Align seg 1/1 to: AAW86355 from: 1 to: 59
3136 ACCCCGCAAGCTCACCCATCTTCTGGCAGTGTCTAAAGAACCCCTG 3183
|||||
28 AsnProGlnAlaHisProTyrPheTrpGlnCysLeuLysAsnAlaLeu 43

seq_name: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:AAW86362
seq_documentation_block:
ID AAW86362 standard; Protein; 100 AA.
XX
XX AAW86362;
XX
XX 15-MAR-1999 (first entry)
XX
XX Mouse DNAX toll-like receptor DTLR6.
XX
XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;
KW modulate inflammatory function; morphological effect;
KW immunological disorder.
XX
XX Mus sp.
XX
XX WO9850547-A2.
XX
XX 12-NOV-1998.
XX
XX 07-MAY-1998; 98WO-US08979.
PF
```

XX 05-MAR-1998; 98US-0076947.  
PR 07-MAY-1997; 97US-0044293.  
PR 22-JAN-1998; 98US-0072212.  
XX (SCHE ) SCHERING CORP.  
XX  
PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;  
XX  
XX WPI; 1999-059670/05.  
DR N-PSDB; AAV80676.  
XX  
XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter  
PT phosphate metabolism, modulate inflammatory function or innate  
PT immunity responses  
XX  
PS Example; Page 150-151; 171pp; English.  
XX  
XX The present invention specifically describes human DNAX toll-like  
CC receptors 2 to 10 (DLR2-10). The present sequence is mouse DTLR6  
CC given in the present invention. Also described are: (1) a fusion  
CC protein comprising a DTLR protein or peptide; (2) a binding compound,  
CC preferably an antibody or antibody fragment which specifically binds to  
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
CC peptide; (4) an expression vector comprising the nucleic acid of (3);  
CC and (5) a host cell comprising the vector of (4). The host cell of (5)  
CC can be used to produce the DTLR proteins. The DTLR proteins can be used  
CC to alter phosphate metabolism, to modulate inflammatory function, innate  
CC immunity responses or morphological effects. The DTLR proteins can be  
CC used in the treatment of conditions exhibiting abnormal expression of  
CC the receptors of their ligands. These abnormalities are typically  
CC manifested by immunological disorders.  
XX  
SQ Sequence 100 AA;  
  
alignment\_scores:  
Quality: 16.00 Length: 16  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x AAW86362 ..  
  
Align seg 1/1 to: AAW86362 from: 1 to: 100  
  
700 TTAAGAGTGTCTCTCCCTGAAGATAACAATGTCACAGCGTCCCTACT 747  
|||||  
15 LeuLysValLeuSerLeuLysAspAsnValThrAlaValProThr 30  
  
seq\_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT: AAG65892  
  
seq\_documentation\_block:  
ID AAG65892 standard; protein; 483 AA.  
XX  
AC AAG65892;  
XX  
DT 11-FEB-2002 (first entry)  
XX  
DE Amino acid sequence of GSK gene Id 90060.  
XX  
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
KW cytostatic; cerebroprotective; vasotropic; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200172961-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 22-MAR-2001; 2001WO-US09226.  
XX

PR 24-MAR-2000; 2000US-192158P.  
PR 28-MAR-2000; 2000US-192688P.  
PR 27-APR-2000; 2000US-200166P.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick RS;  
PI Lai Y;  
XX  
XX WPI: 2001-639223/73.  
DR N-PSDB; AAI67182.  
XX  
XX Isolated polypeptides, which may be peptide hormones, which are  
PT identified by high throughput genome-based biology which identifies  
PT genes and gene products as therapeutic targets for treatment of  
PT diseases such as diabetes and cancer  
XX  
PS Claim 1; Page 65-66; 99pp; English.  
XX  
XX The invention provides polypeptides (AAG65886-65918) which may be peptide  
CC hormones (including insulin, growth hormones, chemokines, cytokines,  
CC neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic  
CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
CC high throughput genome-based biology and polynucleotides (AAI67176-67208)  
CC encoding them. The polypeptides can be expressed by standard recombinant  
CC methodology. The polypeptides are useful in the treatment of disease such  
CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
CC asthma, manic depression, dementia, delirium, mental retardation,  
CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
CC or sexual development disorders, and dysfunctions of the blood cascade  
CC system including those leading to stroke. The polynucleotides may be used  
CC as diagnostic reagents through detecting mutations in the associated gene  
CC and for chromosome localization and for tissue expression studies. The  
CC polypeptides and polynucleotides may also be used as vaccines.  
XX  
SQ Sequence 483 AA;  
  
alignment\_scores:  
Quality: 12.00 Length: 12  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x AAG65892 ..  
  
Align seg 1/1 to: AAG65892 from: 1 to: 483  
  
2194 TTGGACCTCAGCCACCAACCACTGACCACTGTCCCT 2229  
|||||  
173 LeuAspLeuSerHisAsnGlnLeuThrValPro 184  
  
seq\_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT: AAG65893  
  
seq\_documentation\_block:  
ID AAG65893 standard; protein; 605 AA.  
XX  
AC AAG65893;  
XX  
DT 11-FEB-2002 (first entry)  
XX  
DE Amino acid sequence of GSK gene Id 90060.  
XX  
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
KW cytostatic; cerebroprotective; vasotropic; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200172961-A2.  
XX

XX 04-OCT-2001.  
XX 22-MAR-2001; 2001WO-US09226.  
XX 24-MAR-2000; 2000US-192158P.  
XX 28-MAR-2000; 2000US-192668P.  
XX 27-APR-2000; 2000US-200166P.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y;  
PI WPI: 2001-639223/73.  
DR N-PSDB; AAI671183.  
XX Isolated polypeptides, which may be peptide hormones, which are  
PT identified by high throughput genome-based biology which identifies  
PT genes and gene products as therapeutic targets for treatment of  
PT diseases such as diabetes and cancer  
XX Claim 1; Page 67-68; 99pp; English.  
PS The invention provides polypeptides (AAG65886-65918) which may be peptide  
XX hormones (including insulin, growth hormones, chemokines, cytokines,  
CC neuropeptides, integrins, kallikreins, lamins, melanins, natruietic  
CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
CC high throughput genome-based biology and polynucleotides (AAI67176-67208)  
CC encoding them. The polypeptides can be expressed by standard recombinant  
CC methodology. The polypeptides are useful in the treatment of disease such  
CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
CC asthma, manic depression, dementia, delirium, mental retardation,  
CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
CC or sexual development disorders, and dysfunctions of the blood cascade  
CC system including those leading to stroke. The polynucleotides may be used  
CC as diagnostic reagents through detecting mutations in the associated gene  
CC and for chromosome localization and for tissue expression studies. The  
CC polypeptides and polynucleotides may also be used as vaccines.  
XX Sequence 605 AA;  
SQ  
alignment\_scores:                   Quality: 12.00           Length: 12  
                                      Ratio: 1.000           Gaps: 0  
Percent Similarity: 100.000       Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x AAG65893       ..  
Align seg 1/1 to: AAG65893       from: 1 to: 605  
2194 TTGACCTCAGCCACACCAACCTGACCTGTCCT 2229  
295 LeuAspLeuSerHisAsnGlnLeuThrValPro 306  
seq\_name: /SID51/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAE07271  
seq\_documentation\_block:  
ID AAE07271 standard; peptide; 14 AA.  
XX  
AC AAE07271;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human leucine-rich repeat (LRR) signature sequence #6.  
XX  
KW Human; leucine-rich repeat protein-like; LRR; cytostatic; gene therapy;  
KW HIV; Human immunodeficiency Virus; haemophilia; bleeding disorder;

KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW angioplasty-related restenosis; viral infection; rheumatoid arthritis;  
KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX Homo sapiens.  
OS  
XX WO200157261-A1.  
PN  
XX 09-AUG-2001.  
PD  
XX 02-FEB-2001; 2001WO-US03653.  
PF  
XX 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 27-SEP-2000; 2000US-067221.  
XX (HYSE-) HYSEQ INC.  
PA  
XX Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
PI Drmanac RT, Wang M, Chen L, Yang Y;  
PI WPI: 2001-496930/54.  
DR  
XX Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
XX for diagnosing, treating bleeding disorders, myocardial infarction,  
XX atherosclerosis, angioplasty-related restenosis and glomerular diseases  
XX  
PS Claim 12; Page 112; 156pp; English.  
XX The present sequence is human leucine-rich repeat (LRR) signature  
XX sequence which corresponds to residues 252-265 of LRR protein.  
CC LRR protein is involved in protein recognition, cell adhesion,  
CC development, signal transduction, DNA repair, recombination, immune  
CC responses and transcription. LRR DNA and protein are useful for treating,  
CC preventing haemophilia, bleeding disorders (Bernard-Soulier syndrome),  
CC myocardial infarction, thrombosis, atherosclerosis, glomerular diseases,  
CC angioplasty-related restenosis, viral infections, melanomas,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis.  
CC Irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX  
SQ Sequence 14 AA;

alignment\_scores:                   Quality: 11.00           Length: 11  
                                      Ratio: 1.000           Gaps: 0  
Percent Similarity: 100.000       Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x AAE07271       ..  
Align seg 1/1 to: AAE07271       from: 1 to: 14



2185 CTGGAACCTTGGACCTCAGCCACACCAACTG 2217

1 LeuGlThrLeuAspLeuSerHisAsnGlnLeu 11

seq\_name: /SIDS1/gcgdata/geneseq/gene-seq-emb1/AA2001.DAT:AAE06918

seq\_documentation\_block:

ID AAE06918 standard; Protein: 121 AA.

AC AAE06918;

DT 06-NOV-2001 (first entry)

DE Human Garp protein.

Human; leucine-rich repeat protein-like; LRR; cytostatic; gene therapy;  
HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
angioplasty-related restenosis; viral infection; rheumatoid arthritis;  
multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
acute pancreatitis; diabetes mellitus; autoimmune disease; Garp; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 16 /note= "Encoded by TTTC"

FT Misc-difference 27 /note= "Encoded by CTTC"

FT Misc-difference 40 /note= "Encoded by TTTC"

FT Misc-difference 44 /note= "Encoded by GCGTm"

FT Misc-difference 52 /note= "Encoded by GGCCT"

FT Misc-difference 74 /note= "Encoded by CAC"

FT Misc-difference 92 /note= "Encoded by GTG"

FT Misc-difference 93..94 /note= "Encoded by ACGTCTG"

FT Misc-difference 95 /note= "Encoded by TCT"

FT Misc-difference 102 /note= "Encoded by GCGG"

FT Misc-difference 114 /note= "Encoded by GGG"

FT Misc-difference 120 /note= "Encoded by GAG"

FT WO200157261-A1.

PN 09-AUG-2001.

PD 02-FEB-2001; 2001WO-US03653.

PF 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 27-SEP-2000; 2000US-0672221.

XX (HYSE-) HYSEQ INC.

XX Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;

PI Drmanac RT, Wang M, Chen L, Yang Y;

XX WPI; 2001-496930/54.

DR N-PSDB; AAD13549.

XX Novel leucine-rich repeat protein-like polypeptides and polynucleotides

PT for diagnosing, treating bleeding disorders, myocardial infarction,

PT atherosclerosis, angioplasty-related restenosis and glomerular diseases

XX Claim 12; Page 110; 156pp; English.

XX The invention relates to human leucine-rich repeat (LRR) protein-like  
DNA and protein. LRR protein is involved in protein recognition, cell  
adhesion, development, signal transduction, DNA repair, recombination,  
immune responses and transcription. LRR DNA and protein are useful for  
treating, preventing haemophilia, bleeding disorders (Bernard-Soulier  
syndrome), myocardial infarction, thrombosis, atherosclerosis, glomerular  
diseases, angioplasty-related restenosis, viral infections, melanomas,  
immunological disorders (rheumatoid arthritis, multiple sclerosis,  
psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
periodontitis); wound healing, burns, ulcers, incisions and cancer.  
LRR is also useful for proliferation of neural cells and nerve  
regeneration, for treating peripheral nervous system diseases, central  
nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
mechanical and traumatic disorders (spinal cord disorders, head trauma)  
cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
inflammation associated with infection (septic shock, sepsis or systemic  
inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
chronic inflammatory arthritis, pancreatic cell damage from diabetes  
mellitus type 1, graft versus host disease, inflammation associated with  
pulmonary disease, other autoimmune diseases or inflammatory diseases.  
The present sequence is human gap protein used in the exemplification  
of the invention.

SQ Sequence 121 AA;

alignment\_scores:

Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAE06918 ..

Align seg 1/1 to: AAE06918 from: 1 to: 121

2185 CTGGAACCTTGGACCTCAGCCACACCAACTG 2217  
|||||  
91 LeuGlThrLeuAspLeuSerHisAsnGlnLeu 101

seq\_name: /SIDS1/gcgdata/geneseq/gene-seq-emb1/AA1999.DAT:AAW86365

seq\_documentation\_block:

ID AAW86365 standard; Protein: 336 AA.

XX AAW86365;

DT 15-MAR-1999 (first entry)

DE Human DNAX toll-like receptor DTLR10.

XX DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;  
interleukin 1 receptor; phosphate metabolism; innate immunity response;  
modulate inflammatory function; morphological effect;  
immunological disorder.

OS Homo sapiens.

XX WO9850547-A2.

XX 12-NOV-1998.

XX 07-MAY-1998; 98WO-US08979.

XX

```
PR 05-MAR-1998; 98US-0076947.
PR 07-MAY-1997; 97US-0044293.
PR 22-JAN-1998; 98US-0072212.
XX (SCHE ) SCHERING CORP.
XX
XX Bazan JF, Hardiman GT, Kastelein RA, Rock FL;
PI
XX WPI; 1999-059670/05.
DR N-PSDB; AAV80679.
DR
XX Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter
PT phosphate metabolism, modulate inflammatory function or innate
PT immunity responses
XX
XX Claim 9; Page 160-161; 171pp; English.
PS
XX The present invention specifically describes human DNAX toll-like
CC receptors 2 to 10 (DTLR2-10). The present sequence is human DTLR10
CC given in the present invention. Also described are: (1) a fusion
CC protein comprising a DTLR protein or peptide; (2) a binding compound,
CC preferably an antibody or antibody fragment which specifically binds to
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or
CC peptide; (4) an expression vector comprising the nucleic acid of (3);
CC and (5) a host cell comprising the vector of (4). The host cell of (5)
CC can be used to produce the DTLR proteins. The DTLR proteins can be used
CC to alter phosphate metabolism, to modulate inflammatory function, innate
CC immunity responses or morphological effects. The DTLR proteins can be
CC used in the treatment of conditions exhibiting abnormal expression of
CC the receptors of their ligands. These abnormalities are typically
CC manifested by immunological disorders.
XX
SQ Sequence 336 AA;

alignment_scores:
  Quality: 11.00 Length: 11
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAW86365 ..
Align seg 1/1 to: AAW86365 from: 1 to: 336
2860 TTATGTCGAGGAAAGGACTGGTTACCAGGG 2892
|||||
210 LeuCYsLeuGLuGLuArgAspIrpLeuProGly 220

seq_name: /SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAU20527
seq_documentation_block:
ID AAU20527 standard; Protein: 402 AA.
XX
AC AAU20527;
XX
XX 06-DEC-2001 (first entry)
XX
XX Human secreted protein, Seq ID No 519.
XX
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.
XX
OS Homo sapiens.
XX
XX WO200155326-A2.
PN
XX 02-AUG-2001.
PD
```

```
XX 17-JAN-2001; 2001WO-US01347.
PF
XX 31-JAN-2000; 2000US-0179065.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-451931/48.
DR N-PSDB; AAS33236.
DR
XX New nucleic acids and polypeptides, useful for diagnosing, preventing
PT or treating medical conditions
PT
XX Claim 11; SEQ ID No 519; 753pp; English.
PS
XX The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators (e.g.
CC agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/ autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia
CC angina and thrombosis), infections caused by bacteria, viruses and
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),
CC agonists, antagonists and antibodies can also be used to promote wound
CC healing, maintain organs before transplantation, and support cell culture
CC of primary tissues. AAU20342-AAU20666 represent human secreted protein
CC amino acid sequences, and related sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification but was obtained in electronic format directly from WIPO
CC at: ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 402 AA;

alignment_scores:
  Quality: 11.00 Length: 11
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAU20527 ..
Align seg 1/1 to: AAU20527 from: 1 to: 402
2860 TTATGTCGAGGAAAGGACTGGTTACCAGGG 2892
|||||
276 LeuCYsLeuGLuGLuArgAspIrpLeuProGly 286

seq_name: /SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAE07278
seq_documentation_block:
ID AAE07278 standard; Protein: 674 AA.
XX
AC AAE07278;
XX
```

XX DT 06-NOV-2001 (first entry)  
XX DE Human leucine-rich repeat (LRR) protein #1 fragment.  
XX DE  
XX KW Human; leucine-rich repeat protein-like; LRR; cytosstatic; gene therapy;  
KW KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
KW KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW KW angioplasty-related stenosis; viral infection; rheumatoid arthritis;  
KW KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX OS  
XX OS Homo sapiens.  
XX PN WO200157261-A1.  
XX DE  
XX DE 09-AUG-2001.  
XX PF  
XX PF 02-FEB-2001; 2001WO-US03653.  
XX PR 03-FEB-2000; 2000US-0496914.  
XX PR 27-APR-2000; 2000US-0560875.  
XX PR 27-SEP-2000; 2000US-0672221.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
PI Drmanac RT, Wang M, Chen L, Yang Y;  
XX WPI; 2001-496930/54.  
XX PT Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
PT for diagnosing, treating bleeding disorders, myocardial infarction,  
PT atherosclerosis, angioplasty-related stenosis and glomerular diseases  
PT -  
XX PS Claim 12; Page 137-139; 156pp; English.  
XX CC The present sequence is a fragment of human leucine-rich repeat  
CC (LRR) protein. LRR protein is involved in protein recognition, cell  
CC adhesion, development, signal transduction, DNA repair, recombination,  
CC immune responses and transcription. LRR DNA and protein are useful for  
CC treating, preventing haemophilia, bleeding disorders (Bernard-Soulier  
CC syndrome), myocardial infarction, thrombosis, atherosclerosis, glomerular  
CC diseases, angioplasty-related stenosis, viral infections, melanomas,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
CC irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX SQ Sequence 674 AA;

alignment\_scores:  
Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAE07278  
Align seg 1/1 to: AAE07278 from: 1 to: 674  
2185 CTGGAAACTTGGACCTCAGCCACACCACTG 2217  
|||||  
234 LeuGluThrLeuAspLeuSerHisAsnGlnLeu 244  
seq\_name: /SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT.AAE07281  
seq\_documentation\_block:  
ID AAE07281 standard; Protein; 674 AA.  
XX AC  
XX AC AAE07281;  
XX XX  
XX DT 06-NOV-2001 (first entry)  
XX DE  
XX DE Human leucine-rich repeat (LRR) protein #2 fragment.  
XX KW Human; leucine-rich repeat protein-like; LRR; cytosstatic; gene therapy;  
KW KW HIV; Human Immunodeficiency Virus; haemophilia; bleeding disorder;  
KW KW myocardial infarction; thrombosis; atherosclerosis; glomerular disease;  
KW KW angioplasty-related stenosis; viral infection; rheumatoid arthritis;  
KW KW multiple sclerosis; psoriasis; systemic lupus erythematosus; burn; ulcer;  
KW KW inflammatory bowel disease; wound healing; cancer; Alzheimer's disease;  
KW KW Parkinson's disease; cerebrovascular disease; stroke; HIV; fibrosis;  
KW KW lung; liver; nephritis; Crohn's disease; anaphylaxis; hypersensitivity;  
KW KW acute pancreatitis; diabetes mellitus; autoimmune disease.  
XX OS  
XX OS Homo sapiens.  
XX PN WO200157261-A1.  
XX DE  
XX DE 09-AUG-2001.  
XX PF  
XX PF 02-FEB-2001; 2001WO-US03653.  
XX PR 03-FEB-2000; 2000US-0496914.  
XX PR 27-APR-2000; 2000US-0560875.  
XX PR 27-SEP-2000; 2000US-0672221.  
XX XX (HYSE-) HYSEQ INC.  
XX PA Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
XX PI Drmanac RT, Wang M, Chen L, Yang Y;  
XX WPI; 2001-496930/54.  
XX PT Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
PT for diagnosing, treating bleeding disorders, myocardial infarction,  
PT atherosclerosis, angioplasty-related stenosis and glomerular diseases  
PT -  
XX PS Claim 12; Page 151-153; 156pp; English.  
XX CC The present sequence is a fragment of human leucine-rich repeat  
CC (LRR) protein. LRR protein is involved in protein recognition, cell  
CC adhesion, development, signal transduction, DNA repair, recombination,  
CC immune responses and transcription. LRR DNA and protein are useful for  
CC treating, preventing haemophilia, bleeding disorders (Bernard-Soulier  
CC syndrome), myocardial infarction, thrombosis, atherosclerosis, glomerular  
CC diseases, angioplasty-related stenosis, viral infections, melanomas,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis,  
CC irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.

CC Irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock,  
CC chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX  
SQ Sequence 674 AA;

alignment\_scores:  
Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAE07281 ..

Align seg 1/1 to: AAE07281 from: 1 to: 674

2185 CTGGAACTTGGACCTGACCCACCAACCACTG 2217  
|||||  
234 LeuGluThrLeuAspLeuSerHisAsnGlnLeu 244

seq\_name: /SIDSl/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAU29303

seq\_documentation\_block:

ID AAU29303 standard; Protein: 692 AA.

XX  
AC AAU29303;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human PRO polypeptide sequence #280.  
XX  
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;  
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
XX  
OS Homo sapiens.

XX WO200168848-A2.

XX PD 20-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US06520.

XX 01-MAR-2000; 2000WO-US05601.

XX 02-MAR-2000; 2000WO-US05841.

XX 03-MAR-2000; 2000US-187202P.

XX 06-MAR-2000; 2000US-186968P.

XX 14-MAR-2000; 2000US-189320P.

XX 14-MAR-2000; 2000US-189328P.

XX 15-MAR-2000; 2000WO-US06884.

XX 21-MAR-2000; 2000US-190828P.

XX 21-MAR-2000; 2000US-191007P.

XX 21-MAR-2000; 2000US-191048P.

XX 21-MAR-2000; 2000US-191314P.

XX 28-MAR-2000; 2000US-192655P.

XX 29-MAR-2000; 2000US-193032P.

XX 29-MAR-2000; 2000US-193053P.

XX 30-MAR-2000; 2000WO-US08439.

XX 04-APR-2000; 2000US-194449P.

PR 18-APR-2000; 2000US-198585P.  
PR 25-APR-2000; 2000US-199397P.  
PR 25-APR-2000; 2000US-199550P.  
PR 25-APR-2000; 2000US-199654P.  
PR 03-MAY-2000; 2000US-201516P.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 05-JUN-2000; 2000US-209832P.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 22-AUG-2000; 2000US-0644848.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Chen J, Desnovers L, Goddard A, Godowski PJ, Gurney AL;  
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
XX  
DR WPI; 2001-602746/68.  
XX N-PSDB; AAS46204.  
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
PT presence of tumours, such as prostate and breast tumours, in mammals and  
PT to screen for modulators of the compounds -  
XX  
FS Claim 11; Fig 560; 774pp; English.  
XX  
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.  
CC The PRO polypeptides and their associated nucleic acids can be used to  
CC detect the presence of a tumour in a mammal by comparing the level of  
CC expression of a PRO polypeptide in a test sample of cells from the animal  
CC and a control sample of normal cells, whereby a higher level of  
CC expression in the test sample indicates the presence of a tumour in the  
CC animal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats  
CC and rabbits but are preferably human. The polypeptides can be used to  
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,  
CC when contacted with it. A specific polypeptide can be used to stimulate  
CC the proliferation or differentiation of chondrocyte cells. The PRO  
CC proteins can be used to determine the presence of tumours and also  
CC susceptibility to tumour development, particularly adrenal, lung, colon,  
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian  
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids  
CC can be used for genetic analysis of individuals with genetic disorders.  
XX  
SQ Sequence 692 AA;

alignment\_scores:  
Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAU29303 ..

Align seg 1/1 to: AAU29303 from: 1 to: 692

2185 CTGGAACCTTGGACCTGACCCACCAACCACTG 2217

|||||

252 LeuGluThrLeuAspLeuSerHisAsnGlnLeu 262

seq\_name: /SIDSl/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAE07266

seq\_documentation\_block:

ID AAE07266 standard; Protein: 692 AA.

XX AAE07266;

XX DT 06-NOV-2001 (first entry)



FT Region 375..388  
FT /label= Leucine\_rich\_repeat\_signature  
FT Region 378..391  
FT /label= Leucine\_rich\_repeat\_signature  
FT Region 535..548  
FT /label= Leucine\_rich\_repeat\_signature  
FT Region 560..573  
FT /label= Leucine\_rich\_repeat\_signature  
FT Domain 648..673  
FT /label= Transmembrane\_domain  
XX WO200157261-A1.  
PN XX  
XX 09-AUG-2001.  
XX  
XX 02-FEB-2001; 2001WO-US03653.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 27-SEP-2000; 2000US-0672221.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Boyle BJ, Yeung G, Mize NK, Arterburn MC, Tang YT, Liu C;  
PI Drmanac RT, Wang M, Chen L, Yang Y;  
XX  
XX WPI: 2001-496930/54.  
DR N-PSDB; AND13552.  
XX  
XX Novel leucine-rich repeat protein-like polypeptides and polynucleotides  
PT for diagnosing, treating bleeding disorders, myocardial infarction,  
PT atherosclerosis, angioplasty-related restenosis and glomerular diseases  
PT  
XX  
XX Claim 12; Page 149-151; 156pp; English.  
PS  
XX  
XX The present sequence is human leucine-rich repeat (LRR) protein.  
CC LRR protein is involved in protein recognition, cell adhesion,  
CC development, signal transduction, DNA repair, recombination, immune  
CC responses and transcription. LRR DNA and protein are useful for treating,  
CC preventing haemophilia, bleeding disorders (Bernard-Soulier syndrome),  
CC myocardial infarction, thrombosis, atherosclerosis, glomerular diseases,  
CC angioplasty-related restenosis, viral infections, multiple sclerosis,  
CC immunological disorders (rheumatoid arthritis, multiple sclerosis,  
CC psoriasis, systemic lupus erythematosus, inflammatory bowel disease,  
CC periodontitis); wound healing, burns, ulcers, incisions and cancer.  
CC LRR is also useful for proliferation of neural cells and nerve  
CC regeneration, for treating peripheral nervous system diseases, central  
CC nervous system diseases (Alzheimer's, Parkinson's disease, Huntington's  
CC disease, amyotrophic lateral sclerosis and Shy-Drager syndrome);  
CC mechanical and traumatic disorders (spinal cord disorders, head trauma)  
CC cerebrovascular diseases (stroke); HIV, lung or liver fibrosis.  
CC Irritation associated with infection (septic shock, sepsis or systemic  
CC inflammatory response syndrome (SIRS)), ischaemia-reperfusion injury,  
CC endotoxin lethality, arthritis, complement-mediated hyperacute rejection,  
CC nephritis, cytokine or chemokine-induced lung injury, Crohn's disease,  
CC anaphylaxis, hypersensitivity, acute pancreatitis, endotoxin shock.  
CC Chronic inflammatory arthritis, pancreatic cell damage from diabetes  
CC mellitus type 1, graft versus host disease, inflammation associated with  
CC pulmonary disease, other autoimmune diseases or inflammatory diseases.  
XX  
SQ Sequence 692 AA;

alignment\_scores:  
Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Smilarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAE07280  
Align seg 1/1 to: AAE07280 from: 1 to: 692

2185 CTGAAACTTTTGACCTCAGCCACCAACTG 2217  
|||||  
252 LeuGluThrLeuAspLeuSerHisAsnGlnLeu 262  
seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:ABG21179  
seq\_documentation\_block:  
ID ABG21179 standard; Protein; 977 AA.  
XX  
XX AC ABG21179;  
XX  
XX DT 18-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #21170.  
XX  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX  
XX PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX DR WPI: 2001-639362/73.  
DR N-PSDB; AAS85366.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID NO 51538; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 977 AA;

alignment\_scores:  
Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Smilarity: 100.000 Percent Identity: 100.000

```
alignment_block:
US-09-202-054-2 x ABG21179
Align seg 1/1 to: ABG21179 from: 1 to: 977
2860 TTATGCTCGAGGAAGGACTGGTTACCAGG 2892
|||||
829 LeuCysLeuGluGluArgAspTrpLeuProGly 839

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAU29299
seq_documentation_block:
ID AAU29299 standard; Protein; 1032 AA.
XX
AC AAU29299;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #276.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.

XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WL, Zhang Z;
XX
DR WPI; 2001-602746/68.
DR N-PSDB; AAS46200.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
PS Claim 11; Fig 552; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
SQ Sequence 1032 AA;

alignment_scores:
Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAU29299
Align seg 1/1 to: AAU29299 from: 1 to: 1032
2860 TTATGCTCGAGGAAGGACTGGTTACCAGG 2892
|||||
906 LeuCysLeuGluGluArgAspTrpLeuProGly 916

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAU29299
seq_documentation_block:
ID AAU29299 standard; Protein; 1032 AA.
XX
AC AAU29299;
XX
DT 01-OCT-2001 (first entry)
XX
DE Amino acid sequence of a human Toll-like receptor variant TLR9-A.
XX
KW Human; Toll-like receptor; TLR9; vaccine; inflammatory disorder;
KW cardiovascular disorder; systemic infection; autoimmune disease; asthma;
KW rhinitis; chronic obstructive pulmonary disease; emphysema; diabetes;
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease; TLR9-A;
KW rheumatoid arthritis; osteoarthritis; psoriasis; Alzheimer's disease;
KW atherosclerosis; Multiple Sclerosis; septic shock syndrome.
XX
OS Homo sapiens.
XX
PN WO200155386-A1.
XX
PD 02-AUG-2001.
XX
```

```
PF 25-JAN-2001; 2001WO-GB00299.
XX
PR 25-JAN-2000; 2000GB-0001704.
XX
PA (GLAX ) GLAXOSMITHKLINE.
XX
PI Lewis AP, Ray KP;
XX
XX WPI: 2001-457729/49.
DR N-PSDB: AAH42425.
XX
PT An isolated Toll-like receptor polypeptide useful for the treatment or
PT diagnosis of disorders including inflammatory or cardiovascular
PT disorders -
XX
PS Disclosure; Page 37-40; 55pp; English.
XX
CC The present sequence represents human Toll-like receptor (TLR9) variant,
CC designated TLR9-A. The Toll-like receptor protein has immunomodulatory
CC activity, and may be used in vaccines. TLR9 is useful to identify a
CC compound which modulates Toll-like receptor activity. Such compounds
CC are useful to treat an inflammatory or cardiovascular disorder, systemic
CC infection or autoimmune disease that is responsive to Toll-like receptor
CC modulation, for example viral, fungal or bacterial infection, asthma,
CC rhinitis, chronic obstructive pulmonary disease, emphysema, inflammatory
CC bowel disease such as ulcerative colitis or Crohn's disease, rheumatoid
CC arthritis, osteoarthritis, psoriasis, Alzheimer's disease,
CC atherosclerosis, Multiple Sclerosis, diabetes or septic shock syndrome
CC associated with systemic infection involving gram positive or gram
CC negative bacteria. They may also be used to manufacture medicament for
CC the treatment of an immune or inflammatory disorder.
XX
SQ Sequence 1032 AA;

alignment_scores:
  Quality: 11.00 Length: 11
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAG63016 ..

Align seg 1/1 to: AAG63016 from: 1 to: 1032

2860 TTATGCTCGAGGAAAGGACTGGTTACACAGG 2892
|||||
906 LeuCysLeuGluGluArgAspTrpLeuProGly 916

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAG63015
seq_documentation_block:
ID AAG63015 standard; Protein: 1055 AA.
XX
AC AAG63015;
XX
DT 01-OCT-2001 (first entry)
XX
DE Amino acid sequence of a human Toll-like receptor polypeptide.
XX
KW Human; Toll-like receptor; TLR9; vaccine; inflammatory disorder;
KW cardiovascular disorder; systemic infection; autoimmune disease; asthma;
KW rhinitis; chronic obstructive pulmonary disease; emphysema; diabetes;
KW inflammatory bowel disease; ulcerative colitis; Crohn's disease;
KW rheumatoid arthritis; osteoarthritis; psoriasis; Alzheimer's disease;
KW atherosclerosis; Multiple Sclerosis; septic shock syndrome.
XX
OS Homo sapiens.
XX
PN WO200155386-A1.
XX
PD 02-AUG-2001.
XX
```

```
PF 25-JAN-2001; 2001WO-GB00299.
XX
PR 25-JAN-2000; 2000GB-0001704.
XX
PA (GLAX ) GLAXOSMITHKLINE.
XX
PI Lewis AP, Ray KP;
XX
XX WPI: 2001-457729/49.
DR N-PSDB: AAH42424.
XX
PT An isolated Toll-like receptor polypeptide useful for the treatment or
PT diagnosis of disorders including inflammatory or cardiovascular
PT disorders -
XX
PS Claim 1; Page 31-33; 55pp; English.
XX
CC The present sequence represents a human Toll-like receptor polypeptide,
CC designated TLR9. The Toll-like receptor protein has immunomodulatory
CC activity, and may be used in vaccines. TLR9 is useful to identify a
CC compound which modulates Toll-like receptor activity. Such compounds
CC are useful to treat an inflammatory or cardiovascular disorder, systemic
CC infection or autoimmune disease that is responsive to Toll-like receptor
CC modulation, for example viral, fungal or bacterial infection, asthma,
CC rhinitis, chronic obstructive pulmonary disease, emphysema, inflammatory
CC bowel disease such as ulcerative colitis or Crohn's disease, rheumatoid
CC arthritis, osteoarthritis, psoriasis, Alzheimer's disease,
CC atherosclerosis, Multiple Sclerosis, diabetes or septic shock syndrome
CC associated with systemic infection involving gram positive or gram
CC negative bacteria. They may also be used to manufacture medicament for
CC the treatment of an immune or inflammatory disorder.
XX
SQ Sequence 1055 AA;

alignment_scores:
  Quality: 11.00 Length: 11
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAG63015 ..

Align seg 1/1 to: AAG63015 from: 1 to: 1055

2860 TTATGCTCGAGGAAAGGACTGGTTACACAGG 2892
|||||
929 LeuCysLeuGluGluArgAspTrpLeuProGly 939

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAW86356
seq_documentation_block:
ID AAW86356 standard; Protein: 329 AA.
XX
AC AAW86356;
XX
DT 15-MAR-1999 (first entry)
XX
DE Partial human DNAX toll-like receptor DTLR7.
XX
KW DNAX toll-like receptor; DTLR; Drosophila toll receptor; IL-1 receptor;
KW interleukin 1 receptor; phosphate metabolism; innate immunity response;
KW modulate inflammatory function; morphological effect;
KW immunological disorder.
XX
OS Homo sapiens.
XX
PN WO9850547-A2.
XX
PD 12-NOV-1998.
XX
PF 07-MAY-1998; 98WO-US08979.
XX
```



PR 05-MAR-1998; 98US-0076947.  
PR 07-MAY-1997; 97US-0044293.  
PR 22-JAN-1998; 98US-0072212.  
XX  
XX  
PA (SCHE ) SCHERING CORP.  
XX  
XX  
PI Bazan JF, Hardiman GT, Kastelein RA, Rock FL;  
XX  
XX  
DR WPI; 1999-059670/05.  
DR N-PSDB; AAV80670.  
XX  
XX  
PT Human DNAX toll-like receptors, DTLR 2-10 - used to, e.g. alter  
PT phosphate metabolism, modulate inflammatory function or innate  
PT immunity responses  
XX  
XX  
PS Claim 6; Page 131-132; 171pp; English.  
XX  
XX  
CC The present invention specifically describes human DNAX toll-like  
CC receptors 2 to 10 (DTLR2-10). The present sequence is partial human  
CC DTLR7 given in the present invention. Also described are: (1) a fusion  
CC protein comprising a DTLR protein or peptide; (2) a binding compound,  
CC preferably an antibody or antibody fragment which specifically binds to  
CC a DTLR protein or peptide; (3) a nucleic acid encoding a DTLR protein or  
CC peptide; (4) an expression vector comprising the nucleic acid of (3);  
CC and (5) a host cell comprising the vector of (4). The host cell of (5)  
CC can be used to produce the DTLR proteins. The DTLR proteins can be used  
CC to alter phosphate metabolism, to modulate inflammatory function, innate  
CC immunity responses or morphological effects. The DTLR proteins can be  
CC used in the treatment of conditions exhibiting abnormal expression of  
CC the receptors of their ligands. These abnormalities are typically  
CC manifested by immunological disorders.  
XX  
SQ Sequence 329 AA;

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAW86356 ..

Align seg 1/1 to: AAW86356 from: 1 to: 329

844 CTTGACCTAAGTGGAAATGGCCCTGTTCT 873  
|||||  
86 LeuAspLeuSerGlyAsnCysProArgCys 95

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT: AAB64892

seq\_documentation\_block:  
ID AAB64892 standard; Protein: 426 AA.  
XX  
AC AAB64892;  
XX  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 11 SEQ ID NO:70.  
XX  
KW Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;  
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;  
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;  
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vulnary;  
KW antialzheimers; antiparkinsonian; antiparkinsonian; immune disorder;  
KW multiple sclerosis; systemic lupus erythematosus; HIV; infection;  
KW hyperproliferative disorder; cancer; Gaucher's disease; wound healing;  
KW cardiovascular disease; Scimitar syndrome; Chaga's cardiomyopathy;  
KW coronary arteriosclerosis; angiogenic disorder; diabetic retinopathy;  
KW corneal graft neovascularisation; neurological disorder; regeneration;  
KW Huntington's chorea; Alzheimer's disease; Parkinson's disease;  
KW infectious disease; chemotaxis.  
XX

OS Homo sapiens.  
XX  
PN WO2000076530-A1.  
XX  
XX 21-DEC-2000.  
PD  
XX  
XX 01-JUN-2000; 2000WO-US14933.  
XX  
XX 11-JUN-1999; 99US-0138572.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX  
XX  
PI Rosen CA, Ruben SM, Komatsoulis GA;  
XX  
XX  
XX WPI; 2001-071147/08.  
DR N-PSDB; AAF33223.  
XX  
XX  
PT Nucleic acids encoding 49 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
PT disease and diabetic retinopathy -  
XX  
XX  
PS Claim 11; Page 489-490; 554pp; English.  
XX  
XX The polynucleotide sequences given in AAF33213 to AAF33261 encode the  
XX human secreted proteins given in AAB64882 to AAB64930. AAB64931 to  
XX AAB64991 represent human secreted polypeptide sequences and proteins  
XX homologous to them, which are given in the exemplification of the present  
XX invention. Human secreted proteins have activities based on the tissues  
XX and cells the genes are expressed in. Examples of activities include:  
XX immunomodulatory; antisclerotic; dermatological; immunosuppressive;  
XX antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;  
XX vascular; antimicrobial; anti-angiogenic; nootropic; antialzheimers;  
XX neuroprotectant; anticonvulsant; and vulnerary. The polynucleotides and polypeptides can  
XX be used in the prevention, diagnosis and treatment of diseases associated  
XX with inappropriate polypeptide expression. Disorders that may be  
XX prevented, diagnosed and/or treated by the above methods include immune  
XX disorders (e.g. multiple sclerosis, systemic lupus erythematosus and  
XX human immuno-deficiency virus (HIV) infections), hyperproliferative  
XX disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases  
XX (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary  
XX arteriosclerosis), angiogenic disorders (e.g. corneal graft  
XX neovascularisation and diabetic retinopathy), neurological disorders  
XX (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
XX infectious diseases and/or for promoting wound healing, regeneration and  
XX /or chemotaxis. AAF33204 to AAF33212 and AAB64881 represent sequences  
XX used in the exemplification of the present invention.  
XX  
\* SQ Sequence 426 AA;

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAB64892 ..

Align seg 1/1 to: AAB64892 from: 1 to: 426

2185 CTGGAACCTTTGGACCTCAGCCACACCAACAA 2214  
|||||  
252 LeuGluThrLeuAspLeuSerHisnGln 261

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT: AAU14800

seq\_documentation\_block:  
ID AAU14800 standard; Protein: 504 AA.  
XX  
AC AAU14800;  
XX

DT 24-OCT-2001 (first entry)  
DE Novel bone marrow polypeptide #6.  
XX  
XX Bone marrow; immunosuppressive; immunostimulant; cytostatic; vulnary;  
KW nontropic; neuroprotective; therapeutic; antigenic; nutritional source;  
KW cytokine; stem cell growth factor; tissue regeneration; cancer;  
KW Parkinson's disease; Alzheimer's disease; neurodegenerative disorder;  
KW wound healing; immune system; autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
XX WO200155442-A2.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 25-JAN-2001; 2001WO-US02543.  
XX  
XX PR 25-JAN-2000; 2000US-0491404.  
XX PR 17-JUL-2000; 2000US-0617746.  
XX PR 03-AUG-2000; 2000US-0631451.  
XX PR 15-SEP-2000; 2000US-0663870.  
XX PR 30-NOV-2000; 2000US-0250583.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Zhang J;  
PI Dmanac RT;  
XX  
XX DR WPI; 2001-465578/50.  
XX DR N-PSDB; AAS23105.  
XX  
XX PT Novel polynucleotides encoding bone marrow-derived polypeptides useful  
PT for treating, e.g., cancer, autoimmune disease and Alzheimer's disease  
PT  
XX  
XX PS Claim 10; Page 173-174; 274pp; English.  
XX  
XX AAU14795-AAU14973 represent the amino acid sequences of novel bone  
CC marrow-derived polypeptides. The proteins may exhibit e.g., cytokine or  
CC stem cell growth factor activity and may be useful for re-engineering  
CC damaged or diseased tissues, producing large quantities of human cells to  
CC treat Parkinson's, Alzheimer's and other neurodegenerative diseases,  
CC wound healing, immune system stimulation or suppression, treating  
CC autoimmune diseases, and cancer. The corresponding nucleic acid sequences  
CC can be used to express recombinant protein for analysis, characterisation  
CC or therapeutic use; as markers for tissues in which the corresponding  
CC protein is preferentially expressed; as a molecular weight marker on  
CC gels; as chromosome markers or tags; as probes to hybridise and discover  
CC novel, related DNA sequences; as a source of information to derive  
CC polymerase chain reaction (PCR) primers; for selecting and making  
CC oligomers for attachment to a 'gene chip' or other support; to raise  
CC anti-protein antibodies using DNA immunisation techniques; and as an  
CC antigen to raise anti-DNA antibodies or elicit another immune response.  
XX The proteins may be also used as a nutritional source.  
XX  
SQ Sequence 504 AA;

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAU14800 ..  
Align seg 1/1 to: AAU14800 from: 1 to: 504

844 CTTGACCTAAGTGGAAATTCCTCGTTGT 873  
154 LeuAspLeuSerGlyAsnCysProArgCys 163

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:ABB23027  
seq\_documentation\_block:  
ID ABB23027 standard; Protein: 1040 AA.  
XX  
XX AC ABB23027;  
XX  
XX DT 23-JAN-2002 (first entry)  
XX  
XX DE Protein #5026 encoded by probe for measuring heart cell gene expression.  
XX  
XX KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200157274-A2.  
XX  
XX PD 09-AUG-2001.  
XX  
XX PF 30-JAN-2001; 2001WO-US00666.  
XX  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234687.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX DR WPI; 2001-488899/53.  
XX  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
XX PS Claim 15; SEQ ID NO 24797; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA411305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1040 AA;

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x ABB23027 ..  
Align seg 1/1 to: ABB23027 from: 1 to: 1040

844 CTTGACCTAAGTGGAAATTCCTCGTTGT 873  
250 LeuAspLeuSerGlyAsnCysProArgCys 259

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAM31141

```
seq_documentation_block:
ID  AAM31141 standard; Protein; 1040 AA.
XX
AC  AAM31141;
XX
DT  17-OCT-2001 (first entry)
XX
DE  Peptide #5178 encoded by probe for measuring placental gene expression.
XX
KW  Probe; microarray; human; placenta; antenatal diagnosis;
KW  genetic disorder.
XX
OS  Homo sapiens.
XX
PN  WO200157272-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00663.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
PA  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-488897/53.
XX
PT  Human genome-derived single exon nucleic acid probes useful for
PT  analyzing gene expression in human placenta -
XX
PS  Claim 27; SEQ ID No 31410; 654pp; English.
XX
CC  The present invention relates to single exon nucleic acid probes (SENP;
CC  see AA1315-AA157346). The present sequence is a peptide encoded by one
CC  such probe. The probes are useful for producing a microarray for
CC  predicting, measuring and displaying gene expression in samples derived
CC  from human placenta. The probes are useful for antenatal diagnosis of
CC  human genetic disorders.
XX
SQ  Sequence 1040 AA;

alignment_scores:
  Quality: 10.00      Length: 10
  Ratio: 1.000       Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAM31141 ..
Align seg 1/1 to: AAM31141 from: 1 to: 1040
844 CTTGACCTAAGTGGAAATTCCTCGTGT 873
|||||
250 LeuAspLeuSerGlyAsnCysProArgCys 259

seq_name: /SIDSl/cgdata/geneseq/geneseq-emb1/AA1999.DAT.AAY41768

seq_documentation_block:
ID  AAY41768 standard; Protein; 1041 AA.
XX
AC  AAY41768;
XX
DT  07-DEC-1999 (first entry)
XX
DE  Human PRO286 protein sequence.
XX
KW  Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW  probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW  secreted protein; transmembrane protein.
XX
OS  Homo sapiens.
XX
PN  WO9946281-A2.
XX
PD  16-SEP-1999.
XX
PF  08-MAR-1999; 99WO-US05028.
XX
PR  10-MAR-1998; 98US-0077450.
PR  11-MAR-1998; 98US-0077632.
PR  11-MAR-1998; 98US-0077641.
PR  11-MAR-1998; 98US-0077649.
PR  12-MAR-1998; 98US-0077791.
PR  13-MAR-1998; 98US-0078004.
PR  17-MAR-1998; 98US-0040220.
PR  20-MAR-1998; 98US-0078886.
PR  20-MAR-1998; 98US-0078910.
PR  20-MAR-1998; 98US-0078936.
PR  20-MAR-1998; 98US-0078939.
PR  25-MAR-1998; 98US-0079294.
PR  26-MAR-1998; 98US-0079656.
PR  27-MAR-1998; 98US-0079663.
PR  27-MAR-1998; 98US-0079684.
PR  27-MAR-1998; 98US-0079689.
PR  27-MAR-1998; 98US-0079728.
PR  27-MAR-1998; 98US-0079786.
PR  30-MAR-1998; 98US-0079920.
PR  30-MAR-1998; 98US-0079923.
PR  31-MAR-1998; 98US-0080105.
PR  31-MAR-1998; 98US-0080107.
PR  31-MAR-1998; 98US-0080165.
PR  31-MAR-1998; 98US-0080194.
PR  01-APR-1998; 98US-0080327.
PR  01-APR-1998; 98US-0080328.
PR  01-APR-1998; 98US-0080333.
PR  01-APR-1998; 98US-0080334.
PR  08-APR-1998; 98US-0081049.
PR  08-APR-1998; 98US-0081070.
PR  08-APR-1998; 98US-0081071.
PR  09-APR-1998; 98US-0081195.
PR  09-APR-1998; 98US-0081203.
PR  09-APR-1998; 98US-0081229.
PR  15-APR-1998; 98US-0081817.
PR  15-APR-1998; 98US-0081838.
PR  15-APR-1998; 98US-0081952.
PR  15-APR-1998; 98US-0081955.
PR  21-APR-1998; 98US-0082588.
PR  21-APR-1998; 98US-0082589.
PR  22-APR-1998; 98US-0082700.
PR  22-APR-1998; 98US-0082704.
PR  22-APR-1998; 98US-0082804.
PR  23-APR-1998; 98US-0082767.
PR  23-APR-1998; 98US-0082796.
PR  27-APR-1998; 98US-0083336.
PR  28-APR-1998; 98US-0083322.
PR  29-APR-1998; 98US-0083392.
PR  29-APR-1998; 98US-0083495.
PR  29-APR-1998; 98US-0083496.
PR  29-APR-1998; 98US-0083499.
PR  29-APR-1998; 98US-0083500.
PR  29-APR-1998; 98US-0083545.
PR  29-APR-1998; 98US-0083554.
PR  29-APR-1998; 98US-0083558.
PR  29-APR-1998; 98US-0083559.
PR  30-APR-1998; 98US-0083742.
PR  05-MAY-1998; 98US-0084366.
PR  06-MAY-1998; 98US-0084414.
```

```
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
XX (GETH ) GENENTECH INC.
XX
XX Wood WL, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
XX WPI; 1999-551358/46.
XX N-PSDB; AAZ34304.
XX
XX New secreted and transmembrane polypeptides and their polynucleotides,
XX useful for treating blood coagulation disorders, cancers and cellular
XX adhesion disorders -
XX
XX Claim 12; Flg 211; 530pp; English.
XX
XX The present invention describes secreted and transmembrane polypeptides
XX and their polynucleotides. The nucleotide sequences are useful as
XX sources of probes, primers, for chromosome mapping, and for generation
XX of antisense sequences. They can also be used to create transgenic
XX animals. The proteins can be used to treat a variety of diseases and
XX disorders, depending on their function. Diseases that may be treated
XX include blood coagulation disorders, cancers and cellular adhesion
XX disorders. They may also be used to raise antibodies. AAZ33891 to
XX AAZ34338, and AAZ41685 to AAZ41774 represent polynucleotide and
XX polypeptide sequence given in the exemplification of the present
XX invention.
XX
XX Sequence 1041 AA;

alignment_scores:
  Quality: 10.00 Length: 10
  Ratio: 1.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x AAY41768 ..
Align seg 1/1 to: AAY41768 from: 1 to: 1041
844 CTTGACCTAAGTGGAAATGCGCTCGTTCT 873
|||||
251 LeuAspLeuSerGlyAsnGlyProargCys 260

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AAY05867

seq_documentation_block:
ID AAY05867 standard; Protein; 1041 AA.
XX
AC AAY05867;
XX
XX 02-AUG-1999 (first entry)
XX
XX Human Toll protein PRO286.
XX
XX PRO286; Toll; homologue; human; adaptive immunity; septic shock;
XX inflammation; diabetes; amyotrophic lateral sclerosis; cancer;
XX ulcer; rheumatoid arthritis; pathogen pattern recognition receptor;
XX signal transduction.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..26
XX Protein 27..1041
XX Domain 826..848
XX Region 130..151
XX Region 206..227
XX Region 662..684
XX Region 669..690
XX Region 693..714
XX Modified-site 29
XX Modified-site 42
XX Modified-site 80
XX Modified-site 88
XX Modified-site 115
XX Modified-site 160
XX Modified-site 247
XX Modified-site 285
XX Modified-site 293
XX Modified-site 358
XX Modified-site 362
XX Modified-site 395
XX Modified-site 416
XX Modified-site 443
XX Modified-site 511
XX Modified-site 546
XX Modified-site 582
XX Modified-site 590
XX Modified-site 640
XX Modified-site 680
```

FT Modified-site /note= "N-glycosylated"  
FT 752  
FT /note= "N-glycosylated"  
FT 937  
FT Modified-site /note= "N-glycosylated"  
FT 1026  
FT /note= "N-glycosylated"  
XX  
PN W09920756-A2.  
XX  
XX 29-APR-1999.  
XX  
PF 07-OCT-1998; 98WO-US21141.  
XX  
XX 26-JUN-1998; 98US-0105413.  
PR 17-OCT-1997; 97US-0062250.  
PR 13-NOV-1997; 97US-0065311.  
PR 28-APR-1998; 98US-0083322.  
PR 26-JUN-1998; 98US-0090863.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Goddard A, Godowski PJ, Gurney AL, Mark MR, Yang R;  
PI  
XX  
XX WPI; 1999-302739/25.  
DR N-PSDB; AAX58296.  
DR  
XX  
XX New human Toll-like receptors that recognize microbial structures  
PT  
XX  
XX Claim 1; Fig 3; 79pp; English.  
PS  
XX  
XX This is the amino acid sequence of PRO286, a novel human homologue  
CC of Drosophila Toll protein, that acts as a pathogen pattern  
CC recognition receptor, sensing the presence of conserved molecular  
CC structures present on microorganisms. The sequence was deduced  
CC from isolated cDNA clone DNA2663-1154 (ATCC 209386) (see AAX58296).  
CC The invention provides 3 novel cDNA clones that encode novel human  
CC Toll polypeptides PRO285 (see AAY05866), PRO286 and PRO358 (see  
CC AAY05868). It also provides specific antibodies and chimeric  
CC molecules comprising a PRO285, PRO286 or PRO358 polypeptide, or a  
CC transmembrane domain-deleted or inactivated variant, fused to a  
CC heterologous amino acid sequence, such as an epitope tag or  
CC immunoglobulin Fc region. Being homologues of Drosophila Toll  
CC protein, the 3 human proteins are likely to be involved in adaptive  
CC immunity, particularly inflammation, septic shock, and response to  
CC pathogens in diseases aggravated by the immune response, such as  
CC diabetes, amyotrophic lateral sclerosis, cancer, rheumatoid  
CC arthritis and ulcers. The PRO polypeptides are used to identify  
CC other proteins involved in Toll-mediated signal transduction (e.g.  
CC natural ligands), to screen for receptor and ligand mimics, and to  
CC generate antibodies. Antibodies specific for the PRO polypeptides  
CC (or for the known receptor TLR-2) are used to treat septic shock  
CC (claimed).  
XX  
SQ Sequence 1041 AA;

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAY05867 ..

Align seg 1/1 to: AAY05867 from: 1 to: 1041

844 CTTGACCTAAGTGGAAATTGCCCTCGTGTG 873

|||||

251 LeuAspLeuSerGlyAsnCysProArgCys 260

seq\_name: /SIDSI/ycgdata/geneseq/geneseq-emb1/AA2000.DAT.AAB44324

seq\_documentation\_block:  
ID AAB44324 standard; Protein; 1041 AA.  
XX  
AC AAB44324;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human PRO286 protein sequence SEQ ID NO:498.  
XX  
KW Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;  
KW expressed sequence tag; detection; cancer.  
XX  
OS Homo sapiens.  
XX  
XX W0200053756-A2.  
XX  
XX 14-SEP-2000.  
PD  
XX  
PF 18-FEB-2000; 2000WO-US04341.  
XX  
XX 08-MAR-1999; 99WO-US05028.  
PR 12-MAR-1999; 99US-0123957.  
PR 29-MAR-1999; 99US-0126773.  
PR 21-APR-1999; 99US-0130232.  
PR 28-APR-1999; 99US-0131445.  
PR 14-MAY-1999; 99US-0134287.  
PR 23-JUN-1999; 99US-0141037.  
PR 26-JUL-1999; 99US-0145698.  
PR 29-OCT-1999; 99US-0162506.  
PR 30-NOV-1999; 99WO-US28313.  
PR 02-DEC-1999; 99WO-US28551.  
PR 16-DEC-1999; 99WO-US28565.  
PR 30-DEC-1999; 99WO-US30095.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;  
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
PI Kljavin LJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;  
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;  
XX  
XX WPI; 2000-611443/58.  
DR N-PSDB; AAC78584.  
DR  
XX  
XX Novel PRO polypeptides and polynucleotides used in detection methods,  
PT to target bioactive molecules to specific cells, and to modulate  
PT cellular activities -  
XX  
XX Claim 12; Fig 211; 636pp; English.  
PS  
XX  
XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed  
CC sequence tag) sequences which encode secreted or transmembrane PRO  
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic  
CC activity. The polynucleotides and polypeptides can be used for detecting  
CC the presence of PRO polypeptides in samples, for linking bioactive  
CC molecules to cells and for modulating biological activities of cells,  
CC using the polypeptides for specific targeting. The polypeptide targeting  
CC can be used to kill the target cells, e.g. for the treatment of cancers.  
CC The polypeptide pairs provide specific targeting of bioactive molecules  
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in  
CC the isolation of the PRO polynucleotide sequences.  
XX  
SQ Sequence 1041 AA;

alignment\_scores:  
Quality: 10.00 Length: 10

Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAB44324 ..

Align seg 1/1 to: AAB44324 from: 1 to: 1041

844 CTTGACCTAAGTGGAAATTGCCCTCGTTCT 873

251 LeuAspLeuSerGlyAsnCysProArgCys 260

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT:AA13130

seq\_documentation\_block:

ID AAR13130 standard; Protein: 15 AA.

XX AC AAR13130;

XX DT 01-OCT-1991 (first entry)

XX DE GPIb alpha peptide fragment.

XX KW Von Willebrand factor; vWF: platelet membrane glycoprotein Ib;  
glycoalbumin; thrombosis.

XX OS Synthetic.

XX PN W09109614-A.

XX PD 11-JUL-1991.

XX PF 04-JAN-1991; 91WO-US00087.

XX PR 14-NOV-1990; 90US-0613083.

XX PR 04-JAN-1990; 90US-0460674.

XX PA (SCRI-) SCRIPPS CLINIC & RE.

XX PI Ruggeri ZM, Zimmerman TS, Houghten RA, Vicente V, Mohri H;

XX PI Ware JL;

XX DR WPI; 1991-222654/30.

XX PT GPIb alpha peptide fragment - inhibits binding of von Willebrand  
factor to platelet membrane glyco-protein Ib, useful in treating  
thrombosis.

XX PS Claim 1; Page 56; 76pp; English.

XX CC The peptide corresponds to residues 81-95 of the N-terminus of  
glycoalbumin, a water sol. proteolytic fragment of GPIb alpha. It  
may be linked to a second peptide from the 45 kD N-terminal  
cryptic fragment of GPIb alpha. The peptide inhibits binding of  
vWF to GPIb. It can be used to inhibit activation, aggregation  
and/or adhesion of platelets, esp. for inhibition of thrombosis.  
XX CC See also AAR13128-R13138.

XX SQ Sequence 15 AA;

alignment\_scores:

Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAR13130 ..

Align seg 1/1 to: AAR13130 from: 1 to: 15

2191 ACTTGGACCTCAGGCACCAACCACTG 2217

|||||

1 ThrLeuAspLeuSerHisAsnGlnLeu 9

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AA40831

seq\_documentation\_block:

ID AAM40831 standard; Protein: 114 AA.

XX AC AAM40831;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 5762.

XX KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX DR N-PSDB; AAI59987.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders  
such as central nervous system injuries -

XX PS Example 2; SEQ ID NO 5762; 10078pp; English.

XX CC The invention relates to human nucleic acids (AA157798-AA161369) and  
the encoded polypeptides (AAM38642-AA442213) with neurotropic,  
immunosuppressant and cytostatic activity. The polynucleotides are useful  
in gene therapy. A composition containing a polypeptide or polynucleotide  
of the invention may be used to treat diseases of the peripheral nervous  
system, such as peripheral nervous injuries, peripheral neuropathy and  
localised neuropathies and central nervous system diseases, such as  
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
assays for receptor activity, arthritis and inflammation, leukaemias and  
C.N.S disorders.  
XX CC Note: The sequence data for this patent did not form part of the printed  
specification.

XX SQ Sequence 114 AA;

alignment\_scores:

Quality: 9.00 Length: 9

Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x AAM40831 ..  
Align seg 1/1 to: AAM40831 from: 1 to: 114  
2335 TATCTGGATCTCAGCTCAATAAATC 2361  
|||||  
105 TyrLeuAspLeuSerAsnLysIle 113

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAB44116  
seq\_documentation\_block:  
ID AAB44116 standard; Protein: 155 AA.  
AC AAB44116;  
XX 08-FEB-2001 (first entry)  
XX Human cancer associated protein sequence SEQ ID NO:1561.

Human: cancer associated gene; cancer antigen; detection; cancer;  
diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;  
antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;  
antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
dermatological; neuroprotective; thrombolytic; coagulant; nootropic;  
vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
immune disorder; haematopoietic cell disorder; autoimmune disorder;  
allergic reaction; graft versus host disease; organ rejection;  
haemostatic; thrombolytic; cardiovascular disorder; infection;  
neurological disease; drug screening.

OS Homo sapiens.  
XX WO200055350-A1.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US05882.  
XX 12-MAR-1999; 99US-0124270.  
XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;  
XX WPI; 2000-587533/55.  
XX N-PSDB; AAC78325.  
XX Novel isolated nucleic acids comprising sequences encoding peptides  
XX useful for treating or diagnosing e.g. cancer -  
XX Claim 11; Page 2238-2239; 2352pp; English.

XX AAC7607 to AAC78448 encode the human cancer associated proteins given  
XX in AAB43398 to AAB44239. The proteins can have activities based on the  
XX tissues and cells the genes are expressed in. Example of activities  
XX include: cytostatic; proliferative; vulnerable; immunomodulator;  
XX antidiabetic; antiasthmatic; antirheumatic; antiarthritic;  
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
XX nootropic; vasotropic; antipsoriatic and antiangiogenic. The  
XX polynucleotides and polypeptides can be used for preventing, treating or  
XX ameliorating medical conditions and diagnosing pathological conditions.  
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
XX the present invention may be used to treat immune disorders by activating  
XX or inhibiting the proliferation, differentiation or mobilisation of  
XX immune cells, to treat disorders of haematopoietic cells, autoimmune  
XX disorders, allergic reactions, graft versus host disease and organ  
XX rejection, modulate haemostatic or thrombolytic activity, modulate  
XX inflammation, cancers, cardiovascular disorders, neurological disease and

CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
CC the present invention.  
XX Sequence 155 AA;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAB44116 ..  
Align seg 1/1 to: AAB44116 from: 1 to: 155  
508 ATACCGCAGGCTCCCGCTAGCTTA 534  
|||||  
60 IleProGlnGlyLeuProSerLeu 68

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY94900

seq\_documentation\_block:

ID AAY94900 standard; Protein: 158 AA.  
XX AC AAY94900;  
XX 16-JUN-2000 (first entry)

XX Human secreted protein clone ns197\_1 protein sequence SEQ ID NO:6.

XX Human; secreted protein; immunestimulant; immunosuppressant; virucide;  
KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;  
KW antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide;  
KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;  
KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;  
KW connective tissue disease; multiple sclerosis; erythematosis;  
KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;  
KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
KW insulin dependent diabetes mellitus; graft-versus-host-disease;  
KW autoimmune inflammatory eye disease; allergy.

XX Homo sapiens.  
XX WO200009552-A1.  
XX 24-FEB-2000.  
XX 13-AUG-1999; 99WO-US18298.  
XX 14-AUG-1998; 98US-0096622.  
XX 17-AUG-1998; 98US-0096815.  
XX 04-SEP-1998; 98US-0099229.  
XX 23-OCT-1998; 98US-0105368.  
XX 08-JAN-1999; 99US-0115234.  
XX 12-FEB-1999; 99US-0119931.  
XX 18-FEB-1999; 99US-0120575.  
XX 30-APR-1999; 99US-0132020.  
XX 11-AUG-1999; 99US-0096622.  
XX (GEMY ) GENETICS INST INC.

XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;  
PI Wong GG, Clark HF, Fechtel K;  
XX WPI; 2000-205979/18.  
XX New polynucleotides encoding secreted proteins, which may have e.g.  
PT nutritional, chemokine, immune stimulating or suppressing,  
PT hematopoiesis regulating, tissue growth, activin/inhibin

PT antiinflammatory or tumor inhibition activity -  
XX Claim 15; Page 470; 641pp; English.  
XX  
CC AAAL6618 to AAAL6697 encode the human secreted proteins given in  
CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,  
CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,  
CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,  
CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,  
CC and adult bladder, cDNA libraries. The polynucleotides and proteins are  
CC predicted to have biological activities which would make them suitable  
CC for treating, preventing or ameliorating medical conditions in humans  
CC and animals. The polynucleotides can be used as markers for tissues in  
CC which the protein is preferentially expressed, as molecular weight  
CC markers on Southern gels, and as chromosome markers or tags to identify  
CC chromosomes or to map gene positions. The proteins can be used in the  
CC treatment of immune deficiencies and disorders, such as severe combined  
CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other  
CC infections. These infections include human immunodeficiency virus (HIV),  
CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and  
CC candidiasis. The proteins can be used to treat autoimmune disorders such  
CC as connective tissue disease, multiple sclerosis, systemic lupus  
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,  
CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent  
CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and  
CC autoimmune inflammatory eye disease. The proteins can also be used to  
CC treat allergic conditions, such as asthma. AAAL6698 to AAAL6774 represent  
CC probes for the human secreted proteins from the present invention.  
XX  
SQ Sequence 158 AA;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAY94900 ..  
Align seg 1/1 to: AAY94900 from: 1 to: 158  
2335 TATCTGATCTCAGCTCAATAAATC 2361  
|||||  
78 TyrLeuAspLeuSerSerAsnLysIle 86

seq\_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAG24828

seq\_documentation\_block:  
ID AAG24828 standard; Protein; 180 AA.  
XX  
AC AAG24828;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28650.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 18-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.



```
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0146386.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145931.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.

PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x AAG24828 ..
Align seg 1/1 to: AAG24828 from: 1 to: 180

2317 CTTGTAGAAAATACATTCGTGACACTCC 2291
|||||
94 LeuValGluAsnThrSerSerAspSer 102

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAB64951
seq_documentation_block:
ID AAB64951 standard; Protein: 227 AA.
XX
AC AAB64951;
XX
DT 23-MAR-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 11 SEQ ID NO:129.
XX
KW Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vulneryary;
KW antiAlzheimers; antiparkinsonian; antimicrobial; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV; infection;
KW hyperproliferative disorder; cancer; Gaucher's disease; wound healing;
KW cardiovascular disease; Scimitar syndrome; Chaga's cardiomyopathy;
KW coronary arteriosclerosis; angiogenic disorder; diabetic retinopathy;
KW corneal graft neovascularisation; neurological disorder; regeneration;
KW Huntington's chorea; Alzheimer's disease; Parkinson's disease;
KW infectious disease; chemotaxis.
XX
OS Homo sapiens.
XX
PN WO200076530-A1.
XX
```

PD 21-DEC-2000.  
XX  
XX  
XX 01-JUN-2000; 2000WO-US14933.  
XX  
XX 11-JUN-1999; 99US-0138572.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX  
XX Rosen CA, Ruben SM, Komatsoulis GA;  
XX WPI; 2001-071147/08.  
XX  
XX Nucleic acids encoding 49 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
PT disease and diabetic retinopathy -  
XX  
XX Disclosure: Page 527-528; 554pp; English.  
XX  
XX The polynucleotide sequences given in AAF33213 to AAF33261 encode the  
CC human secreted proteins given in AAB64882 to AAB64930. AAB64931 to  
CC AAB64991 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Examples of activities include:  
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;  
CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;  
CC vascular; antimicrobial; anti-angiogenic; ophthalmological;  
CC neuroprotectant; anticonvulsant; nootropic; antiaizheimers;  
CC antiparkinsonian; and vulnerary. The polynucleotides and polypeptides can  
CC be used in the prevention, diagnosis and treatment of diseases associated  
CC with inappropriate polypeptide expression. Disorders that may be  
CC prevented, diagnosed and/or treated by the above methods include immune  
CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and  
CC human immuno-deficiency virus (HIV) infections), hyperproliferative  
CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases  
CC (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary  
CC arteriosclerosis), angiogenic disorders (e.g. corneal graft  
CC neovascularisation and diabetic retinopathy), neurological disorders  
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
CC infectious diseases and/or for promoting wound healing, regeneration and  
CC /or chemotaxis. AAF3204 to AAF33212 and AAB64881 represent sequences  
XX used in the exemplification of the present invention.  
XX  
SQ Sequence 227 AA:

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAB64951 ..

Align seg 1/1 to: AAB64951 from: 1 to: 227

2185 CTGGAACCTTGACCTCAGCCACAC 2211  
219 LeuGlutThrLeuAspLeuSerHisAsn 227

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT:AA42265

seq\_documentation\_block:  
ID AAR42265 standard; Protein: 234 AA.  
XX  
AC AAR42265;  
XX  
DT 28-APR-1994 (first entry)  
XX  
DE Decorin sequence PT-76 (N-terminal to LRR8).  
XX  
KW leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;

KW fusion protein; maltose binding protein; tumour growth; inhibition;  
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.  
XX  
PN W09320202-A.  
XX  
PD 14-OCT-1993.  
XX  
PF 02-APR-1993; 93WO-US03171.  
XX  
PR 03-APR-1992; 92US-0865652.  
XX  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
XX  
XX Cardenas J, Craig W, Mullen DG, Pierschbacher MD;  
PI Ruoslahti EI;  
XX  
XX WPI; 1993-336910/42.  
DR N-PSDB; AAQ50051.  
XX  
XX Active fragments of protein esp. decorin - with cell regulatory  
PT factor domain, useful for inhibiting cell regulatory factor  
PT activity  
XX  
PS Claim 10; Page 45-46; 77pp; English.  
XX  
XX Active fragments of decorin (full-length coding sequence AAQ50046)  
CC were generated by PCR and fused to Maltose Binding Protein. The  
CC resulting fusion proteins were useful for inhibiting the activity of  
CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
CC conditions associated with over-activity of the growth factor such  
CC as certain tumours.  
XX  
SQ Sequence 234 AA:

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAR42265 ..

Align seg 1/1 to: AAR42265 from: 1 to: 234

508 ATACCGCAGGGCTCCCGCTAGCTTA 534  
187 IleProGlnGlyLeuProSerLeu 195

seq\_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAU20516

seq\_documentation\_block:  
ID AAU20516 standard; Protein: 235 AA.  
XX  
AC AAU20516;  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Human secreted protein, Seq ID No 508.  
XX  
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytostatic; Alzheimer's disease; Parkinson's disease; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing.  
XX  
OS Homo sapiens.  
XX  
XX W0200155326-A2.  
XX  
XX 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01347.  
XX PR 31-JAN-2000; 2000US-0179065.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX DR WPI; 2001-451931/48.  
XX DR N-PSDB; AAS33225.  
XX  
XX PT New nucleic acids and polypeptides, useful for diagnosing, preventing  
XX PT or treating medical conditions -  
XX PS Claim 11; SEQ ID No 508; 753pp; English.  
XX  
XX CC The invention relates to novel isolated nucleic acid molecules (I)  
XX CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
XX CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
XX CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
XX CC the prevention, treatment and diagnosis of diseases associated with  
XX CC inappropriate expression of secreted proteins. (I) and complementary  
XX CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
XX CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
XX CC of similar nucleic acid sequences in samples, and so which patients may  
XX CC be in need of restorative therapy. (II) may also be used as antigens in  
XX CC the production of antibodies and in assays to identify modulators  
XX CC (agonists and antagonists) of the expression and activity of the secreted  
XX CC proteins. The anti-(II) antibodies and antagonists may also be used to  
XX CC down regulate expression and activity of (II). The anti-(II) antibodies  
XX CC may also be used as diagnostic agents for detecting the presence of (II)  
XX CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The  
XX CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
XX CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
XX CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
XX CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
XX CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
XX CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
XX CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,  
XX CC angina and thrombosis), infections caused by bacteria, viruses and  
XX CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),  
XX CC agonists, antagonists and antibodies can also be used to promote wound  
XX CC healing, maintain organs before transplantation, and support cell culture  
XX CC of primary tissues. AAU20342-AAU20666 represent human secreted protein  
XX CC amino acid sequences, and related sequences of the invention.  
XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC specification but was obtained in electronic format directly from WIPO  
XX CC at: ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 235 AA;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAU20516 ..  
Align seg 1/1 to: AAU20516 from: 1 to: 235

2335 TATCTGGATCTCAGCTCAATAATAATC 2361  
|||||  
82 TyrLeuAspLeuSer-SerAsnIysIle 90

seq\_name: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1993.DAT.AAR42266

seq\_documentation\_block:  
ID AAR42266 standard; Protein: 280 AA.  
XX  
AC AAR42266;

XX 28-APR-1994 (first entry)  
XX DT Decorin sequence PT-77 (N-terminal to LRR10).  
XX DE  
XX KW leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;  
KW fusion protein; maltose binding protein; tumour growth; inhibition;  
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.  
XX  
XX PN WO9320202-A.  
XX PD 14-OCT-1993.  
XX PF 02-APR-1993; 93WO-US03171.  
XX PR 03-APR-1992; 92US-0865652.  
XX PA (LJOL-) LA JOLLA CANCER RES FOUND.  
XX PI Cardenas J, Craig W, Mullen DG, Pierschbacher MD;  
PI Ruoslahti EI;  
XX WPI: 1993-336910/42.  
XX DR N-PSDB; AAQ50052.  
XX  
XX PT Active fragments of protein esp. decorin - with cell regulatory  
XX PT factor domain, useful for inhibiting cell regulatory factor  
XX PT activity  
XX PS Claim 10; Page 47-48; 77pp; English.  
XX  
XX CC Active fragments of decorin (full-length coding sequence AAQ50046)  
XX CC were generated by PCR and fused to Maltose Binding Protein. The  
XX CC resulting fusion proteins were useful for inhibiting the activity of  
XX CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
XX CC conditions associated with over-activity of the growth factor such  
XX CC as certain tumours.  
XX SQ Sequence 280 AA;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x AAR42266 ..  
Align seg 1/1 to: AAR42266 from: 1 to: 280

508 ATACGCGAGGCGCTCCGCTAGCTTA 534  
|||||  
187 IleProGlnGlyLeuProSerLeu 195

seq\_name: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1989.DAT.AAP91368

seq\_documentation\_block:  
ID AAP91368 standard; peptide; 293 AA.  
XX  
XX AC AAP91368;  
XX  
XX DT 21-MAR-1990 (first entry)  
XX DE 45 kDa amino terminal tryptic fragment of glyocalicin and derivs.  
XX KW Glyocalicin; von Willebrand factor; platelet membrane glycoprotein Ib;  
KW platelet aggregation prevention; thrombosis inhibition; antithrombotic  
KW agent  
XX  
XX OS Homo sapiens.  
XX  
XX FH Key Location/Qualifiers

FT Peptide 18...34  
FT Peptide 21...35  
FT Peptide 26...40  
FT Peptide 26...34  
FT Peptide 141...155  
FT Peptide 231...245  
FT Peptide 271...285  
FT Peptide 281...285  
XX EP317278-A.  
PN  
XX  
XX 24-MAY-1989.  
PD  
XX  
PF 16-NOV-1988; 88EP-0310799.  
XX  
XX 17-NOV-1987; 87US-0121454.  
PR  
XX  
XX (SCRI-) SCRIPPS CLINIC RES.  
PA  
XX Zimmerman TS, Ruggeri ZM, Houghten RA, Vincete V, Mohri H;  
PI  
XX WPI; 1989-152756/21.  
DR  
XX  
XX Proteolytic 45 KD fragment of glycoalbumin and derivs. - which inhibit  
PT binding of von Willebrand factor to platelet membrane glyco:protein, and  
PT used as antithrombotic agents  
PT  
XX  
XX Claim 1; : 10pp; English.  
PS  
XX  
XX The claim is for a peptide of a 45kd amino terminal tryptic fragment of  
CC glycoalbumin selected from the sequence which inhibits binding of von  
CC Willebrand factor to platelet membrane glycoprotein 1b and related  
CC molecules or other cells and cell matrices. Also claimed are a sequential  
CC subset of the above (Claim 2) and specific peptides (see FT) (Claim 3)  
CC with the same functions and any peptide of any sequential subset of amino  
CC acids of the sequence (Claim 4). The peptides and derivs. prevent  
CC platelet aggregation and inhibit thrombosis.  
XX  
XX  
XX Sequence 293 AA;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAP91368 ..

Align seg 1/1 to: AAP91368 from: 1 to: 293

2191 ACTTTGACCTCAGCCACACCAACTG 2217

|||||

81 ThrLeuAspLeuSerHisAsnGlnLeu 89

seq\_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:AA42267

seq\_documentation\_block:

ID AAR42267 standard; Protein; 305 AA.

XX AAR42267;

XX 28-APR-1994 (first entry)

XX Decorin sequence PT-78 (N-terminal to half C-terminal).

XX leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;  
KW fusion protein; maltose binding protein; tumour growth; inhibition;  
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.

XX WO9320202-A.

PN

XX 14-OCT-1993.

PD

XX 02-APR-1993; 93WO-US03171.  
PF  
XX 03-APR-1992; 92US-0865652.  
PR  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
PA  
XX Cardenas J, Craig W, Mullen DG, Pierschbacher MD;  
PI Ruoslahti EI;  
PI  
XX WPI; 1993-336910/42.  
DR  
XX N-PSDB; AAQ50053.  
DR  
XX Active fragments of protein esp. decorin - with cell regulatory  
PT factor domain, useful for inhibiting cell regulatory factor  
PT activity  
PT  
XX  
XX Claim 10; Page 49-50; 77pp; English.  
PS  
XX Active fragments of decorin (full-length coding sequence AAQ50046)  
CC were generated by PCR and fused to Maltose Binding Protein. The  
CC resulting fusion proteins were useful for inhibiting the activity of  
CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
CC conditions associated with over-activity of the growth factor such  
CC as certain tumours.  
XX  
XX Sequence 305 AA;

alignment\_scores:

Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AAR42267 ..

Align seg 1/1 to: AAR42267 from: 1 to: 305

508 ATACCGCAGGCCTCCGCCTAGCTTA 534

|||||

187 IleProGlnGlyLeuProSerLeu 195

seq\_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:AA42260

seq\_documentation\_block:

ID AAR42260 standard; Protein; 331 AA.

XX AAR42260;

XX 28-APR-1994 (first entry)

XX Mature decorin PT-65.

XX leucine-rich repeat; proteoglycan; cell regulatory factor; MBP;  
KW fusion protein; maltose binding protein; tumour growth; inhibition;  
KW decorin; PG-II; PG-40; transforming growth factor-beta; TGF-beta.

XX Key Location/Qualifiers

FT Region 1..45

FT /label= N-terminal\_region

FT /note= "contains 4 Cys residues"

FT Region 46..280

FT /label= repeat\_region

FT /note= "contains 10 leucine-rich repeats"

FT Region 281..331

FT /label= C-terminal\_region

XX WO9320202-A.

PN

XX 14-OCT-1993.

XX

XX 02-APR-1993; 93WO-US03171.

PF

XX 03-APR-1992; 92US-0865652.  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
XX Cardenas J, Craig W, Mullen DG, Pierschbacher MD;  
PI Ruoslahti EI;  
XX WPI: 1993-336910/42.  
DR N-PSDB; AAQ50046.  
XX Active fragments of protein esp. decorin - with cell regulatory  
PT factor domain, useful for inhibiting cell regulatory factor  
PT activity  
XX Claim 10; Page 36-38; 77pp; English.  
XX Active fragments of decorin (full-length coding sequence AAQ50046)  
CC were generated by PCR and fused to Maltose Binding Protein. The  
CC resulting fusion proteins were useful for inhibiting the activity of  
CC a cell regulatory factor, esp. TGF-beta, and hence for treating  
CC conditions associated with over-activity of the growth factor such  
CC as certain tumours.  
XX Sequence 331 AA;  
SQ  
alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x AAR42260 ..  
Align seg 1/1 to: AAR42260 from: 1 to: 331  
508 ATACCGCAGGCCTCCCGCTAGCTTA 534  
|||||  
187 IleProGlnGlyLeuProSerLeu 195  
seq\_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:AAR89439  
seq\_documentation\_block:  
ID AAR89439 standard; Protein; 342 AA.  
XX AC AAR89439;  
XX DT 20-AUG-1996 (first entry)  
XX Human recombinant decorin.  
XX Decorin; PG-II; PG-40; proteoglycan; guanidinium ion.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..14  
FT /label= Sig\_peptide  
XX WO9601842-A1.  
XX PD 25-JAN-1996.  
XX PF 07-JUL-1995; 95WO-US08542.  
XX PR 08-JUL-1994; 94US-0272919.  
XX (LJOL-) LA JOLLA CANCER RES FOUND.  
XX Craig WS, Harper JR, Hernandez SD, Kostel PJ, Parker JR;  
PI Vedvick TS;  
XX

DR WPI: 1996-097586/10.  
DR N-PSDB; AAT10741.  
XX Purifich. of human recombinant decorin - using a strong anion  
PT exchange resin, a hydrophobic interaction chromatography resin and a  
PT strong anion exchange resin  
XX Disclosure; Fig 1A-D; 55pp; English.  
XX Human recombinant decorin (AAR89439) was obtd. by expression of a  
CC cDNA clone (AAT10741) in CHO host cells. Decorin (or PGII or PG-40)  
CC is a proteoglycan having a 40 kda core protein. Recombinant  
CC decorin can be produced by cotransfection of CHO-DG44 cells with  
CC pSV2-decorin and pSV2dhfr. Large-scale cultures can be performed  
CC using CHO cells attached to microcarrier beads. The recombinant  
CC protein is purified from the cells using a 3-step chromatographic  
CC procedure. It can be used for the highly sensitive detection of  
CC guanidinium ions (ppm range), partic. in protein-contg. soins.  
XX purified using GuHCl, and also has therapeutic applns.  
SQ Sequence 342 AA;  
alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x AAR89439 ..  
Align seg 1/1 to: AAR89439 from: 1 to: 342  
508 ATACCGCAGGCCTCCCGCTAGCTTA 534  
|||||  
199 IleProGlnGlyLeuProSerLeu 207

---

OM of: US-09-202-054-2 to: PIR\_71:\* out\_format : pfs

Date: Jul 16, 2002 6:22 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+\_n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09202054/runat\_16072002\_074913\_7757/app\_query.fasta\_1.3393  
-DB=PIR\_71 -OPMT=fastan -SUFFIX=oli6.rpr -GAPOP=4\_500  
-GAPEXT=0\_050 -MINMATCH=0\_100 -LOOPEXT=0\_000  
-GAPOP=4\_500 -GAPEXT=0\_050 -GAPOP=60\_000 -LOOPEXT=60\_000  
-GAPEXT=6\_000 -GAPEXT=7\_000 -GAPOP=60\_000 -YGAPEXT=60\_000  
-DELOP=6\_000 -DELEX=7\_000 -START=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR\_SCORE=quality  
-THR\_MIN=6 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORN=ext  
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09202054\_@CGN1\_1\_87 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-202-054-2

Query length: 3283

Database: PIR\_71:\*

Database sequences: 283138

Database length: 96089334

Search time (sec): 108.030000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000

WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
PIR2:T03322	+	12.00	182.25	0.0038	1232	! hypothetical protein F18F4_240
PIR2:AC1328	+	10.00	153.35	0.5146	367	! internalin proteins homolog lmc
PIR2:S67265	+	10.00	147.90	0.4803	791	! hypothetical protein YOR353c -
PIR2:C96673	+	9.00	134.99	5.72	348	! gamma-tocopherol methyltransfer
PIR1:NBHUC8	+	9.00	134.78	5.71	358	! hypothetical protein F7H2.5 - A
PIR2:S06280	+	9.00	134.76	5.71	359	! decorin precursor - human
PIR2:I47020	+	9.00	134.74	5.71	360	! decorin precursor - bovine
PIR1:NBHUIA	+	9.00	130.82	5.43	626	! platelet glycoprotein Ib alpha
PIR2:A49674	+	9.00	126.61	5.15	1134	! chaoptin precursor - fruit fly
PIR2:AS2321	+	9.00	126.81	5.10	1268	! flightless-I homolog - human
PIR2:S56221	+	9.00	123.44	4.94	1770	! hypothetical protein YFL033c -
PIR2:C96615	+	9.00	123.39	4.94	1784	! hypothetical protein F18I24_10
PIR2:T21312	-	9.00	121.63	4.83	2287	! hypothetical protein F23D12.2
PIR2:S68976	-	8.00	126.16	71.84	86	! tetrahydromethanopterin S-methyl
PIR2:B69021	-	8.00	126.16	71.84	86	! tetrahydromethanopterin S-methyl
PIR2:C70010	+	8.00	124.22	70.10	113	! Na+/H+ antiporter homolog yufv
PIR2:A95285	+	8.00	122.50	68.58	144	! hypothetical protein SMA0343 [i
PIR2:C97180	-	8.00	122.26	68.37	149	! uncharacterized protein yih2 fa
PIR2:A03864	+	8.00	121.41	67.64	168	! hypothetical 17.7K protein - hu
PIR2:D83323	+	8.00	120.69	67.02	186	! probable acetyltransferase PA25
PIR2:T02792	+	8.00	120.57	66.93	189	! hypothetical protein L549.4 [im
PIR2:A13239	+	8.00	120.17	66.59	200	! conserved hypothetical protein
PIR1:S73914	+	8.00	120.10	66.53	202	! MG105 homolog K04_orf202 - Mycc
PIR2:T01040	+	8.00	120.03	66.47	204	! hypothetical protein YUP8H12R.2
PIR2:F70469	-	8.00	119.40	65.94	223	! enolase-phosphatase E-1 - Aquif
PIR2:T37467	+	8.00	118.97	65.58	237	! ribosomal protein s4e - Thermop
PIR2:S05572	+	8.00	118.94	65.55	238	! hypothetical protein 238 - Stre
PIR2:S10005	+	8.00	117.71	64.54	288	! hypothetical protein 5 - fowl a
PIR2:A10010	+	8.00	117.58	64.44	288	! pyridoxal kinase (EC 2.7.1.35)
PIR2:C83972	-	8.00	117.11	64.05	308	! 2-dehydropantoate 2-reductase a
PIR2:F86295	+	8.00	116.99	63.96	313	! T24B18.18 protein - Arabidopsis
PIR2:T49908	+	8.00	115.79	62.99	371	! hypothetical protein T24H18.110
PIR2:D71175	+	8.00	115.40	62.68	392	! hypothetical protein P06000 - P
PIR2:H96536	+	8.00	114.26	61.78	460	! hypothetical protein F2J10_8 [i
PIR2:C64119	+	8.00	114.02	61.59	476	! ADPglucose--starch(bacterial g)
PIR2:T51702	+	8.00	113.58	61.25	506	! amidophosphoribosyltransferase
PIR2:B69205	-	8.00	113.47	61.17	514	! cobyric acid synthase - Methano

PIR2:T32012	-	8.00	113.28	61.02	528	! hypothetical protein K02F6.7
PIR2:T00824	+	8.00	113.19	60.95	535	! probable thioredoxin reducta
PIR2:S26857	+	8.00	113.16	60.93	537	! isocitrate lyase (EC 4.1.3.1)
PIR2:A44287	+	8.00	112.84	60.68	562	! pyruvate kinase, plastid - c
PIR2:E96598	+	8.00	112.29	60.26	607	! protein K2N2.2 [imported] -
PIR2:T10727	+	8.00	112.23	60.22	612	! protein kinase xaz1 (EC 2.7.
PIR2:S74727	+	8.00	112.06	60.08	627	! precorrin methylease (EC 2.1.
PIR2:S47299	+	8.00	111.96	60.01	636	! gene F protein - rinderpest
PIR2:T77895	+	8.00	111.83	59.91	648	! hypothetical protein T4C21.2
PIR2:G89894	-	8.00	111.66	59.78	664	! protein kinase [imported] -
PIR2:S04128	+	8.00	111.18	59.42	710	! phenylalanine ammonia-lyase
PIR2:C84633	+	8.00	110.86	59.17	743	! probable disease resistance
PIR2:T17462	+	8.00	110.62	59.00	768	! disease resistance E - tomat
PIR2:T45899	+	8.00	110.49	58.90	783	! receptor protein kinase-like
PIR2:E84514	-	8.00	110.43	58.85	789	! hypothetical protein At2g141
PIR2:A53256	+	8.00	110.28	58.74	806	! nuclear protein b1m4 - Emeri
PIR2:A96557	+	8.00	110.00	58.54	838	! probable receptor protein ki
PIR1:GNLJGB	-	8.00	109.89	58.45	852	! pol polyprotein - bovine leu
PIR1:GNLJGA	-	8.00	109.89	58.45	852	! pol polyprotein - bovine leu
PIR2:S29358	-	8.00	109.88	58.44	853	! disease resistance protein D
PIR2:T17461	+	8.00	109.86	58.43	855	! Cf-4A protein - tomato
PIR2:T07015	+	8.00	109.86	58.43	855	! disease resistance protein -
PIR2:T17460	+	8.00	109.86	58.38	863	! Cf-9 protein precursor - tom
PIR2:A55173	+	8.00	109.42	58.10	910	! hypothetical protein F1017.1
PIR2:B96770	+	8.00	108.89	57.71	981	! chemotactic transducer-relat
PIR2:H82403	+	8.00	108.87	57.70	983	! conserved large membrane pro
PIR2:A87063	-	8.00	108.84	57.68	987	! receptor protein kinase homo
PIR2:T50850	+	8.00	108.74	57.60	1002	! probable transmembrane prot
PIR2:T00876	-	8.00	108.69	57.57	1008	! transport protein tpi - Myc
PIR2:S72698	-	8.00	108.58	57.49	1025	! protein kinase Xaz1 (EC 2.7.
PIR1:A57676	+	8.00	108.46	57.40	1042	! isoleucine--tRNA ligase (EC
PIR2:H70203	+	8.00	107.95	57.03	1120	! hypothetical protein F14D7.
PIR2:B86479	+	8.00	107.80	56.93	1143	! hypothetical protein T13K14
PIR2:T10636	+	8.00	107.75	56.89	1152	! hypothetical protein F13K14
PIR2:AE1852	+	8.00	107.65	56.82	1161	! DNA polymerase III, alpha c
PIR2:B90387	-	8.00	107.69	56.82	1167	! hypothetical protein At2g18
PIR2:E96963	+	8.00	107.53	56.73	1187	! autoantigen - human
PIR2:C84568	+	8.00	107.30	56.57	1215	! protoporphyrin IX magnesium
PIR2:I52882	+	8.00	106.84	56.24	1309	! hypothetical protein SPAC2G
PIR2:H64479	+	8.00	106.79	56.20	1318	! hypothetical protein SPAC2G
PIR2:S62457	+	8.00	106.73	56.16	1329	! WD-repeat containing protei
PIR2:AE1901	-	8.00	106.67	56.11	1341	! hypothetical protein DKF2p4
PIR2:T17285	+	8.00	106.36	55.89	1401	! cyclic peptide synthetase -
PIR1:EDBEGA	+	8.00	106.29	55.84	1415	! immediate-early protein IE1
PIR2:S54587	+	8.00	106.20	55.78	1433	! CAT8 protein - yeast (Sacch
PIR2:B86398	+	8.00	105.46	55.26	1590	! protein T7N9.24 [imported]
PIR2:T31095	+	8.00	104.63	54.68	1788	! vitellogenin precursor - Or
PIR2:SA6622	-	8.00	103.78	54.10	2014	! probable membrane protein y
PIR2:A86216	+	8.00	101.81	52.76	2658	! protein T23G18.2 [imported]
PIR2:C86007	+	7.00	114.21	866.68	33	! hypothetical protein 24773 [i
PIR2:F71704	+	7.00	112.67	849.93	41	! ribosomal protein L36 - Ricke
PIR2:G87660	+	7.00	112.67	849.93	41	! ribosomal protein L36 [import
PIR2:S28769	+	7.00	112.67	849.93	41	! hypothetical protein 41 [psbi
PIR2:S28769	+	7.00	112.67	849.93	41	! 50S ribosomal protein L36 [im
PIR2:B97785	+	7.00	109.40	815.42	65	! cytochrome P450 (clone 5) - M
PIR2:S35170	+	7.00	109.08	812.12	68	! tms protein - Bacillus megate
PIR2:S18901	+	7.00	108.48	805.96	74	! hypothetical protein YN9466.0
PIR2:S50923	-	7.00	108.39	804.99	75	! hypothetical protein Atu4672
PIR2:AD3131	-	7.00	108.11	802.16	78	! ubiquinol--cytochrome-c reduc
PIR1:CC8011	-	7.00	107.33	794.32	87	! NADH dehydrogenase (ubiquinol
PIR2:AG0276	+	7.00	107.01	791.11	91	! Ig kappa chain V region - mou
PIR2:B71391	+	7.00	107.01	791.11	91	! Ig kappa chain V region - mou
PIR2:G90732	-	7.00	107.01	791.11	91	! hypothetical protein ECS0831
PIR2:H91035	+	7.00	107.01	791.11	91	! hypothetical protein ECS3256
PIR2:E65011	+	7.00	107.01	791.11	91	! hypothetical protein b2376 -
PIR2:AB5583	-	7.00	107.01	791.11	91	! unknown protein encoded by pr
PIR2:B85860	+	7.00	107.01	791.11	91	! hypothetical protein 23641 [i
PIR2:AH1065	+	7.00	107.01	791.11	91	! hypothetical protein STY4854
PIR2:S17625	+	7.00	106.86	789.57	93	! Ig kappa chain V region - mou
PIR2:B87025	-	7.00	106.78	788.81	94	! hypothetical protein [importe
PIR2:S78736	-	7.00	106.63	787.32	96	! protein YOL013W-a - yeast (Sa

pir2:PH1062	-	7.00	106.49	785.86	98	Ig light chain V region (clone 2	pir2:A84769	+	7.00	102.71	749.07	167	hypothetical protein At2g354
pir2:S28134	-	7.00	106.49	785.86	98	gas vesicle protein gvpJ [import	pir2:S26718	+	7.00	102.58	747.87	170	hypothetical 19.8K protein (
pir2:E72664	+	7.00	106.13	782.35	103	hypothetical protein APE0740 -	pir2:C84492	+	7.00	102.58	747.87	170	hypothetical protein At2g105
pir2:S50936	-	7.00	106.07	781.67	104	probable membrane protein YDR02	pir2:B81828	-	7.00	102.50	747.09	172	hypothetical integral membra
pir2:E64150	-	7.00	105.86	779.67	107	hypothetical protein H10376 - E	pir2:T49816	+	7.00	102.29	745.16	177	hypothetical protein B24H17
pir2:F79826	-	7.00	105.86	779.67	107	hypothetical protein RC1014 [im	pir2:T27193	+	7.00	102.25	744.79	178	hypothetical protein Y57A108
pir2:T49016	+	7.00	105.86	777.10	107	hypothetical protein FC22.30 -	pir2:T33522	+	7.00	102.25	744.79	178	hypothetical protein T10B11
pir2:S69536	+	7.00	105.60	777.10	111	hypothetical protein 28 - phage	pir2:C81342	+	7.00	102.21	744.41	179	probable 16S rRNA processing
pir2:AD0406	+	7.00	105.48	775.86	113	arsenical resistance operon rep	pir2:F64455	+	7.00	102.17	744.04	180	hypothetical protein homolog
pir2:T48112	+	7.00	105.48	775.86	113	hypothetical protein F18X2.80	pir2:A31788	-	7.00	102.17	744.04	180	signal peptidase (EC 3.4.99.
pir2:B71088	+	7.00	105.48	775.86	113	hypothetical protein PH0967 - E	pir2:T46017	+	7.00	102.13	743.67	181	hypothetical protein T10K17.
pir2:S28127	-	7.00	105.41	775.24	114	gas-vesicle protein gvpJ - Halo	pir2:E70229	-	7.00	101.98	742.21	185	conserved hypothetical prote
pir2:F71123	-	7.00	105.41	775.24	114	gas-vesicle protein gvpJ - Halo	pir2:A84375	-	7.00	101.94	741.85	186	transcription initiation fac
pir2:T08235	+	7.00	105.41	775.24	114	probable gas-vesicle protein gvp	pir2:H70238	+	7.00	101.87	741.13	188	conserved hypothetical protein
pir2:A23329	+	7.00	105.35	774.63	115	Ly-5-8 glycoprotein - mouse (fr	pir2:T06555	+	7.00	101.83	740.78	189	blue copper-binding protein
pir1:BV6CAR	+	7.00	105.23	773.43	117	arsenical resistance operon rep	pir2:T31525	+	7.00	101.83	740.78	189	hypothetical protein Y43F11A
pir1:YQ6CB2	-	7.00	105.11	772.25	119	fimbrial protein precursor - Es	pir2:B72330	+	7.00	101.79	740.43	190	transcription regulator, Tet
pir2:I40630	-	7.00	105.11	772.25	119	fimbrial protein precursor - Es	pir2:B72330	+	7.00	101.79	740.43	190	transcription regulator, Tet
pir2:D71473	+	7.00	105.11	772.25	119	prepilin - Escherichia coli pla	pir2:C69502	-	7.00	101.75	740.08	191	hypothetical protein AFE2020
pir2:B72067	+	7.00	105.11	772.25	119	probable lojap - Chlamydia trac	pir2:H72767	+	7.00	101.75	740.08	191	hypothetical protein APE0131
pir2:B86557	+	7.00	105.11	772.25	119	conserved hypothetical protein	pir2:AE2559	-	7.00	101.75	740.08	191	hypothetical protein al18068
pir1:YQ6BR9	+	7.00	105.05	771.67	120	CT814.1 hypothetical protein [i	pir2:T21210	+	7.00	101.72	739.73	192	hypothetical protein F21F12.
pir2:S53206	-	7.00	105.05	771.67	120	fimbrial protein precursor - Es	pir2:S67699	+	7.00	101.68	739.39	193	probable membrane protein YD
pir2:F71126	+	7.00	105.05	771.67	120	hypothetical protein YJR018w -	pir2:SA5627	+	7.00	101.61	738.70	195	acidic endoprotease precu
pir2:E83836	+	7.00	105.05	771.67	120	hypothetical protein PH0780 - E	pir2:A65105	-	7.00	101.57	738.36	196	phosphoheptose isomerase hom
pir1:YQ6CFX	-	7.00	104.99	771.10	121	hypothetical protein BH1413 [im	pir2:F91132	-	7.00	101.57	738.36	196	hypothetical protein ECS4030
pir1:YQ6CFX	-	7.00	104.99	771.10	121	fimbrial protein precursor - Es	pir2:A85978	-	7.00	101.57	738.36	196	hypothetical protein yraO [i
pir2:S34233	+	7.00	104.99	771.10	121	hypothetical protein, 13K - Pla	pir2:AG0431	-	7.00	101.57	738.36	196	probable phosphosugar isom
pir2:AG2304	+	7.00	104.88	769.96	123	hypothetical protein alr3990 [i	pir2:AF0899	-	7.00	101.57	738.36	196	probable phosphoheptose isom
pir2:S64446	+	7.00	104.82	769.96	124	probable membrane protein vxr13	pir2:H81681	+	7.00	101.57	738.36	196	maf protein TC0628 [importe
pir2:S31006	+	7.00	104.76	768.84	125	gene 61 protein - Mycobacterium	pir2:I50507	+	7.00	101.57	738.36	196	achaete-scute homolog a - ze
pir2:B64922	+	7.00	104.76	768.84	125	probable membrane protein bl648	pir2:C84951	+	7.00	101.53	738.02	197	alkyl hydroperoxide reductas
pir2:E90923	+	7.00	104.76	768.84	125	hypothetical protein ECs2358 [i	pir2:E72743	+	7.00	101.53	738.02	197	hypothetical protein APE0476
pir2:A85772	+	7.00	104.76	768.84	125	hypothetical protein Z2655 [im	pir2:S55886	-	7.00	101.53	738.02	197	CCNH finger protein 6 - Arab
pir2:S49604	+	7.00	104.76	768.84	125	hypothetical protein 126 (rps12	pir2:C53293	-	7.00	101.53	738.02	197	flm3 region hypothetical pro
pir2:D96518	+	7.00	104.70	768.29	127	protein T2B6.14 [imported] - Ar	pir2:AB3212	+	7.00	101.53	738.02	197	2-hydroxychromene-2-carboxyl
pir2:F72517	+	7.00	104.65	767.75	127	hypothetical protein APE2116 -	pir2:B72866	+	7.00	101.50	737.69	198	viral capsid protein - Autog
pir2:S67869	+	7.00	104.43	765.61	131	probable membrane protein YDR05	pir2:A97404	-	7.00	101.50	737.69	198	hypothetical protein AGR_C_6
pir2:S77948	+	7.00	104.32	764.57	133	major allergen Par j I (clone F	pir2:A13621	-	7.00	101.46	737.35	199	conserved hypothetical prote
pir2:B82768	+	7.00	104.27	764.05	134	virulence regulator XF0749 [im	pir2:T32132	+	7.00	101.39	736.69	201	hypothetical protein SC10A7.
pir2:T09650	-	7.00	104.11	762.53	137	ribosomal protein L16 - Vigna	pir2:T23192	+	7.00	101.36	736.36	202	hypothetical protein K01G12.
pir2:AB3544	-	7.00	104.11	762.53	137	transcription regulator alr7637	pir2:AG1280	+	7.00	101.36	736.36	202	1-acylglycerol-3-phosphate O
pir2:AD3568	-	7.00	104.11	762.53	137	integral membrane protein [impc	pir2:AG1643	+	7.00	101.36	736.36	202	1-acylglycerol-3-phosphate O
pir2:D72727	+	7.00	104.06	762.03	138	hypothetical protein APE0357 -	pir2:G81700	+	7.00	101.32	736.04	203	thymidylate kinase TC0460 [i
pir2:D88955	-	7.00	104.01	761.54	139	protein K04F1.9 [imported] - Ca	pir2:E71700	+	7.00	101.32	736.04	203	hypothetical protein RP422
pir2:T10257	-	7.00	103.91	760.56	141	2S albumin precursor - cucurbit	pir2:I54685	+	7.00	101.25	735.39	205	perA protein - Escherichia c
pir2:T40635	-	7.00	103.81	759.60	143	clathrin coat assembly protein	pir2:D97698	+	7.00	101.25	735.39	205	glucose inhibited division p
pir2:AB5072	+	7.00	103.81	759.60	143	hypothetical protein AT4907390	pir2:AE2924	+	7.00	101.25	735.39	205	glucose inhibited division p
pir2:E71222	+	7.00	103.76	759.12	144	hypothetical protein PH0040 - E	pir2:T33097	+	7.00	101.22	735.06	206	cytochrome c oxidase chain I
pir2:E83286	+	7.00	103.76	759.12	144	conserved hypothetical protein	pir2:AE1925	+	7.00	101.18	734.74	207	cytochrome c oxidase chain I
pir2:H82498	+	7.00	103.71	758.65	145	hypothetical protein VCA0109 [i	pir2:D65507	+	7.00	101.05	733.48	211	hypothetical protein T1C2C2.
pir2:JC4691	+	7.00	103.66	758.18	146	coagulation factor IX/factor X	pir2:DA8952	+	7.00	101.01	733.17	212	hypothetical protein YHR110w
pir2:H73201	+	7.00	103.66	758.18	146	hypothetical protein PA80088 -	pir2:D64442	+	7.00	100.95	732.55	214	conserved hypothetical prote
pir2:A71115	+	7.00	103.66	758.18	146	hypothetical protein PA80088 -	pir2:S39831	+	7.00	100.95	732.55	214	hypothetical protein YBL057C
pir2:LEJH30	+	7.00	103.56	757.25	148	lysosome (EC 3.2.1.17) c precu	pir4:S58320	+	7.00	100.95	732.55	214	hypothetical protein orf 610
pir2:T20053	+	7.00	103.56	757.25	148	lysosome (EC 3.2.1.17) c precu	pir2:B35534	+	7.00	100.91	732.24	215	hypothetical 23K protein - L
pir2:T82735	+	7.00	103.56	757.25	148	ribosomal protein Ddl27a - slim	pir2:T18842	+	7.00	100.91	732.24	215	hypothetical protein C01G6.2
pir2:T82735	+	7.00	103.56	757.25	148	hypothetical protein APE0418 -	pir2:AB0080	+	7.00	100.88	731.94	216	probable membrane proteiny
pir2:H90762	+	7.00	103.42	755.89	151	hypothetical protein ECS1072 [i	pir2:C86350	+	7.00	100.85	731.63	217	protein F8K7.12 [imported] -
pir2:T07758	+	7.00	103.37	755.44	152	disease resistance protein homc	pir2:F72214	+	7.00	100.81	731.33	218	hypothetical protein - therm
pir2:S67057	-	7.00	103.28	754.55	154	hypothetical protein YOR169c -	pir2:T34523	+	7.00	100.78	731.03	219	hypothetical protein DKP2p56
pir2:AB6788	+	7.00	103.19	753.68	156	transcription regulator [import	pir2:T49456	+	7.00	100.69	730.14	222	hypothetical protein sl11318
pir2:S52414	+	7.00	103.14	753.24	157	ubiquitin--protein ligase (EC 6	pir2:G87533	+	7.00	100.69	730.14	222	hypothetical protein CC2296
pir2:T31645	+	7.00	103.14	752.81	158	hypothetical protein Y57A10A.w	pir2:S78154	+	7.00	100.69	730.14	222	DNA-directed RNA polymerase
pir2:T43454	+	7.00	103.05	752.39	159	hypothetical protein [imported]	pir2:S02216	+	7.00	100.62	729.55	224	deoxyribose-phosphate aldola
pir2:S19163	+	7.00	103.05	752.39	159	homeotic protein distal-less -	pir2:S48467	+	7.00	100.62	729.55	224	probable membrane protein YL
pir2:S28434	+	7.00	103.01	751.96	160	interferon-induced protein, 15K	pir2:B72527	+	7.00	100.62	729.55	224	hypothetical protein APE2191
pir2:D86418	-	7.00	102.97	751.54	161	unknown protein, 69696-70273 [i	pir2:T49446	+	7.00	100.59	729.25	225	hypothetical protein B17C10.
pir2:E81263	-	7.00	102.92	751.12	162	probable periplasmic thiredoxin	pir2:A33937	+	7.00	100.53	728.67	227	Iq light chain (I301) - horn
pir2:D97787	+	7.00	102.92	751.12	162	hypothetical protein RC0700 [im	pir2:D84058	+	7.00	100.53	728.67	227	hypothetical protein BH3268
pir2:T15255	+	7.00	102.88	750.71	163	hypothetical protein K07B1.6 -	pir2:H83001	+	7.00	100.43	727.81	230	probable permease of ABC tra
pir2:F71691	+	7.00	102.79	749.88	165	hypothetical protein RP348 - R1	pir2:D86352	+	7.00	100.43	727.81	230	protein T26F1.14 [imported]
pir2:D81049	-	7.00	102.71	749.07	167	hypothetical protein NMB1733 [i	pir2:H81322	+	7.00	100.37	727.25	232	probable permease of ABC tra
							pir2:B96908	-	7.00	100.37	727.25	232	(FS) similar to ABC transpor



pir2:T45359	7.00	100.34	726.97	233	two-component regulator [import	pir1:H69215	-	7.00	98.89	713.69	286	agmatine ureohydrolase - Met
pir2:S09760	-	7.00	100.31	234	hypothetical protein TR111 pred	pir2:T48572	+	7.00	98.89	713.69	286	hypothetical protein T31B5.7
pir2:D07083	-	7.00	100.25	236	probable two component system x	pir2:F65165	+	7.00	98.86	713.47	287	33.2 kD protein in dind-rph
pir2:E86635	-	7.00	100.25	236	hypothetical protein yaiH [impc	pir2:G91193	+	7.00	98.86	713.47	287	probable alpha helix protein
pir1:D69136	-	7.00	100.22	237	anaerobic ribonucleotide reduct	pir2:H86040	+	7.00	98.86	713.47	287	probable alpha helix protein
pir2:S03090	+	7.00	100.22	237	probable membrane protein YMR13	pir2:AF0970	+	7.00	98.86	713.47	287	conserved hypothetical prote
pir2:AE2037	+	7.00	100.22	237	hypothetical protein Alr1851 [i	pir2:F72307	+	7.00	98.86	713.47	287	conserved hypothetical prote
pir2:T47954	+	7.00	100.10	241	exonuclease RRP41 [imported] -	pir2:T37029	-	7.00	98.84	713.24	288	hypothetical protein SCJ12.1
pir2:S63458	+	7.00	100.10	241	hypothetical protein YPL024w -	pir2:JC5843	+	7.00	98.81	713.02	289	chitinase (EC 3.2.1.14) III
pir2:T14791	+	7.00	100.07	242	hypothetical protein DKFZp586E0	pir2:G70336	-	7.00	98.81	713.02	289	hypothetical protein aq.406
pir2:T14428	-	7.00	100.04	243	thiamatin-like protein - turnip	pir2:I67465	-	7.00	98.79	712.80	290	arylamine N-acetyltransferase
pir2:C72776	-	7.00	100.04	243	hypothetical protein APE0197 -	pir2:T31470	+	7.00	98.79	712.80	290	hypothetical protein R08A2.5
pir2:T51481	+	7.00	100.02	244	hypothetical protein T2I1H19.30	pir2:T08189	+	7.00	98.77	712.58	291	hypothetical protein T2284.8
pir2:H86197	+	7.00	100.02	244	hypothetical protein [imported]	pir2:T49172	+	7.00	98.72	712.14	293	hypothetical protein T20N10.
pir2:T04700	+	7.00	100.02	244	hypothetical protein T19K4.30 -	pir2:AG0913	+	7.00	98.69	711.92	294	probable adenine-specific DN
pir1:A64579	-	7.00	99.95	246	molybdenum ABC transporter, per	pir2:C83127	-	7.00	98.64	711.49	296	probable ABC transporter regul
pir2:E71934	-	7.00	99.96	246	molybdate ABC transporter, per	pir2:C95410	+	7.00	98.64	711.49	296	probable ABC transporter, pe
pir2:F89262	+	7.00	99.96	246	protein C34D1.4 [imported] - Ca	pir2:T33780	+	7.00	98.57	710.84	299	hypothetical protein C39F7.1
pir2:T84210	-	7.00	99.93	247	oxidoreductase homolog [importe	pir2:H83666	+	7.00	98.53	710.42	301	protein F26F24.12 [imported] -
pir2:T17677	-	7.00	99.93	247	hypothetical protein Al87L - CH	pir2:F86440	+	7.00	98.53	710.42	301	unknown protein [imported] -
pir2:T32514	-	7.00	99.93	248	hypothetical protein C44B12.1 -	pir2:B75162	+	7.00	98.50	710.20	302	ribosome abc transporter, perm
pir2:D83140	+	7.00	99.90	248	hypothetical protein PA4049 [im	pir2:B71179	+	7.00	98.50	710.20	302	hypothetical protein PH1711
pir2:T37920	+	7.00	98.87	249	hypothetical protein SPAC1866.0	pir2:E91252	+	7.00	98.48	709.99	303	tail fiber [imported] - Esch
pir2:T18662	+	7.00	99.84	250	hypothetical protein C02D4.1 -	pir2:AC0392	+	7.00	98.48	709.99	303	probable LysK-family transfr
pir2:G97565	+	7.00	99.84	250	hypothetical protein AGR_C_3137	pir2:T42566	+	7.00	98.45	709.78	304	hypothetical protein Alr8532
pir2:AE3786	-	7.00	99.84	250	conserved hypothetical protein	pir2:T06607	+	7.00	98.41	709.36	306	hypothetical protein F15J13.
pir2:AG1022	-	7.00	99.76	253	probable AraC family regulatory	pir2:T16457	+	7.00	98.39	709.16	307	hypothetical protein F55D10.
pir2:T41192	-	7.00	99.73	254	outer membrane protein 2 - Esch	pir2:B97002	+	7.00	98.36	708.95	308	uncharacterized conserved me
pir2:T33401	+	7.00	99.70	255	hypothetical protein B0545.4 -	pir2:C64370	+	7.00	98.32	708.54	310	modification methylase [EC 2
pir2:AI2344	-	7.00	99.67	256	two-component system response x	pir2:G67671	-	7.00	98.32	708.54	310	probable MAP kinase T9L24.32
pir2:T84946	+	7.00	99.67	256	hypothetical protein yafI [impc	pir2:T86425	+	7.00	98.29	708.33	311	unknown protein [imported] -
pir2:T46548	+	7.00	99.65	257	probable transcription activato	pir2:AF0144	+	7.00	98.29	708.33	311	penicillin-binding protein 7 -
pir1:J16865	+	7.00	99.65	257	gene F41 protein - mouse	pir2:AD1352	-	7.00	98.25	707.92	313	S. aureus Cbf1 protein homol
pir1:Y0635	-	7.00	99.56	260	rRNA (adenine-N6-)-methyltransf	pir2:AG1722	-	7.00	98.25	707.92	313	S. aureus Cbf1 protein homol
pir2:T05536	-	7.00	99.56	260	acid phosphatase (EC 3.1.3.2) -	pir2:JC4951	-	7.00	98.23	707.72	314	troponin T - scallop (Chlamy
pir2:B70419	+	7.00	99.56	260	hypothetical protein aq.1369 -	pir2:AF1695	+	7.00	98.20	707.52	315	transcription repressor of d
pir2:T36700	+	7.00	99.54	261	branched-chain alpha keto acid	pir2:AF3191	+	7.00	98.18	707.31	316	transcription regulator, Lys
pir2:C83566	+	7.00	99.51	262	hypothetical protein PA0642 [im	pir2:T39869	+	7.00	98.16	707.11	317	probable lysophospholipase (
pir2:G03476	+	7.00	99.51	262	lysosomal-associated multitrans	pir2:T37336	+	7.00	98.16	707.11	317	MG302 homolog A05_orf317 - M
pir2:F86278	+	7.00	99.48	263	hypothetical protein F14L17.18	pir2:A95282	-	7.00	98.16	707.11	317	ABC transporter, permease SM
pir2:AG3222	-	7.00	99.48	263	nitrate transport ATP-binding p	pir2:B64900	+	7.00	98.14	706.91	318	hypothetical protein bl471 -
pir1:F68100	-	7.00	99.46	264	conserved hypothetical protein	pir2:CE9192	+	7.00	98.14	706.91	318	stomatol-like protein - Meth
pir2:H98259	+	7.00	99.46	264	conserved hypothetical protein	pir2:S62295	+	7.00	98.14	706.91	318	hypothetical protein - Sync
pir2:E98124	+	7.00	99.46	264	hypothetical protein ABC-MSP [i	pir2:A27741	+	7.00	98.11	706.71	319	rRNA (adenine-N6-)-methyltra
pir2:C83090	+	7.00	99.43	265	probable permease of ABC transp	pir2:S46824	+	7.00	98.09	706.52	320	ribose-phosphate pyrophospha
pir2:T29973	+	7.00	99.43	265	hypothetical protein ZK682.7 -	pir2:SA3453	+	7.00	98.05	706.12	322	CCCl protein - yeast (Sacccha
pir2:F95252	+	7.00	99.40	266	PTS system, IID component [impc	pir2:F84616	+	7.00	98.05	706.12	322	probable bHLH transcription
pir2:D98117	+	7.00	99.40	266	hypothetical protein PTS-EII [i	pir2:G64949	+	7.00	98.02	705.92	323	probable S-adenosylmethionin
pir2:S02045	+	7.00	99.40	266	chlorophyll a/b-binding protein	pir2:E69051	+	7.00	98.02	705.92	323	probable enzyme [imported] -
pir2:T75557	-	7.00	99.40	266	conserved hypothetical protein	pir2:A85800	+	7.00	98.02	705.92	323	probable enzyme yecp [import
pir2:T37123	+	7.00	99.35	268	probable zinc-binding oxidoredu	pir2:AF0744	+	7.00	98.02	705.92	323	conserved hypothetical prote
pir2:T46902	+	7.00	99.35	268	hypothetical protein - Deinococ	pir2:H70073	-	7.00	97.98	705.53	325	two-component sensor histidi
pir2:C75283	+	7.00	99.30	270	hypothetical protein - Deinococ	pir2:T13020	+	7.00	97.96	705.34	326	peroxidase (EC 1.1.1.17) ATP
pir2:T18915	-	7.00	99.24	272	hypothetical protein C04F12.8 -	pir2:CE4480	-	7.00	97.96	705.34	326	DNA repair protein RAD2 homo
pir2:T03254	+	7.00	99.22	273	probable carbonate dehydratase	pir2:T51811	+	7.00	97.96	705.34	326	protein farnesyltransferase
pir2:T52385	+	7.00	99.22	273	zinc finger protein 2, C2H2-type	pir1:DEFWHA	+	7.00	97.94	705.14	327	malate dehydrogenase (EC 1.1
pir2:T04257	-	7.00	99.22	274	hypothetical protein F20B18.130	pir2:AE0296	+	7.00	97.94	705.14	327	phenylalanine-tRNA ligase (
pir2:B97711	-	7.00	99.19	274	hypothetical protein kdsA [impc	pir2:T00797	+	7.00	97.94	705.14	327	hypothetical protein At2g327
pir2:F71714	-	7.00	99.17	275	2-dehydro-3-deoxyphosphoactona	pir2:F64537	-	7.00	97.92	704.95	328	A/G-specific adenine glycosy
pir2:T42826	-	7.00	99.17	275	hypothetical protein C04F12.8 -	pir2:E71969	-	7.00	97.92	704.95	328	a/g-specific adenine glycosy
pir2:A31972	+	7.00	99.17	275	lectin DB58 precursor - horse g	pir2:F89914	-	7.00	97.92	704.95	328	hypothetical protein opp-2B
pir2:T05853	+	7.00	99.14	276	phytohemagglutinin L precursor	pir1:B70177	+	7.00	97.87	704.56	330	conserved hypothetical prote
pir2:S08554	+	7.00	99.11	277	hypothetical protein F27B13.170	pir1:JW0097	+	7.00	97.81	703.99	333	bicoid-related homeobox prot
pir2:B71522	+	7.00	99.09	278	probable deoxyheptonate aldolas	pir2:AG0183	-	7.00	97.77	703.61	335	galactose-binding protein [i
pir2:S52582	+	7.00	99.06	279	prephenate dehydratase (EC 4.2.	pir2:H87247	+	7.00	97.77	703.61	335	probable cell division prote
pir2:AC2264	-	7.00	99.06	279	urease accessory protein D [imf	pir2:A87300	-	7.00	97.74	703.42	336	TPR domain protein [imported
pir2:T01924	+	7.00	99.06	279	hypothetical protein F2F3.15 -	pir2:A64303	+	7.00	97.70	703.05	338	conserved hypothetical prote
pir2:S75831	+	7.00	99.04	280	hypothetical protein slr1530 -	pir1:S04898	+	7.00	97.66	702.67	340	myb-related protein 1 - maiz
pir2:H89800	+	7.00	99.04	280	hypothetical protein SA0339 [im	pir2:T49006	-	7.00	97.66	702.67	340	farnesyltransferase subunit
pir2:C83692	+	7.00	98.96	283	hypothetical protein BH0339 [im	pir2:T49429	-	7.00	97.66	702.67	340	conserved hypothetical prote
pir2:F71015	-	7.00	98.94	284	hypothetical protein PH1420 - F	pir1:K1BE36	+	7.00	97.64	702.49	341	thymidine kinase (EC 2.7.1.2
pir2:E84616	-	7.00	98.94	284	probable bHLH transcription fac	pir2:B53125	+	7.00	97.64	702.49	341	restriction enzyme BcgI beta
pir2:T48440	-	7.00	98.91	285	hypothetical protein T32M21.40	pir2:H81317	+	7.00	97.62	702.30	342	probable lipopolysaccharide

pir2:AB3116	-	7.00	97.62	702.30	342	transcription regulator, LacI 4	pir2:T04322	+	7.00	96.74	694.54	387	polygalacturonase (EC 3.2.1.1
pir2:T42129	-	7.00	97.60	702.12	343	probable acyltransferase (EC 2.3.1.1)	pir2:JN0793	+	7.00	96.74	694.54	387	adaptive-response sensory-k1
pir2:C84607	-	7.00	97.58	701.93	344	hypothetical protein At2g21960	pir2:T204581	-	7.00	96.74	694.54	387	hypothetical protein T0608.8
pir2:A75576	+	7.00	97.58	701.93	344	oxidoreductase - Deinoxococcus ra	pir2:T23452	+	7.00	96.74	694.54	387	hypothetical protein B0412.1
pir2:T40367	+	7.00	97.58	701.93	344	hypothetical protein SPBC3084.0	pir2:T882178	+	7.00	96.74	694.54	387	conserved hypothetical proteo
pir2:D84012	+	7.00	97.56	701.75	345	N-acetylglutamate gamma-semiald	pir2:T33653	+	7.00	96.74	694.54	387	probable DNA repair and reco
pir2:AB3526	+	7.00	97.56	701.75	345	hypothetical protein PA0946 [im	pir2:AG0835	+	7.00	96.74	694.54	387	probable type I secretion pr
pir2:AB2134	-	7.00	97.52	701.39	347	transcription regulator alr2625	pir2:S53975	+	7.00	96.71	694.22	389	probable membrane protein YM
pir2:A46030	+	7.00	97.52	701.39	347	probable membrane protein YBR15	pir2:B64740	-	7.00	96.67	693.90	391	yaeg protein - Escherichia c
pir2:H97720	+	7.00	97.50	701.20	348	hypothetical protein RC0162 [im	pir2:F85500	-	7.00	96.67	693.90	391	hypothetical protein yaeg [l
pir2:T30093	-	7.00	97.48	701.02	348	hypothetical protein C0603.5 -	pir2:F89774	+	7.00	96.67	693.74	392	phosphopentomutase [importe
pir2:F84246	-	7.00	97.48	701.02	349	hypothetical protein Vng0903c	pir2:T15703	-	7.00	96.63	693.58	393	hypothetical protein C29H12.
pir2:H98303	+	7.00	97.43	700.66	351	hypothetical protein AGR_L_2777	pir2:T30313	+	7.00	96.63	693.42	394	hypothetical protein SSO1550
pir2:AE2979	+	7.00	97.43	700.66	351	aldo/keto reductase mocA [import	pir2:S47987	+	7.00	96.60	693.26	395	actin-related protein ARP14D
pir2:D0264	+	7.00	97.43	700.66	351	biotin synthase (bioB) [importe	pir2:E88533	+	7.00	96.56	692.95	397	probable zinc finger protein
pir2:C98320	-	7.00	97.43	700.66	351	hypothetical protein AGR_L_3023	pir2:A38926	+	7.00	96.54	692.79	398	aspartyl proteinase SAP3 (EC
pir2:AB2963	-	7.00	97.43	700.66	351	endoglucanase [imported] - Agri	pir2:T84972	+	7.00	96.53	692.63	399	nicotinate phosphoribosyltra
pir2:E98170	+	7.00	97.39	700.31	353	transcription regulator (ABO114	pir2:C83953	-	7.00	96.53	692.63	399	tyrosyl-tRNA synthetase 2 PA
pir2:AH3116	+	7.00	97.39	700.31	353	transcriptional regulator, LacI 4	pir2:D70072	+	7.00	96.53	692.63	399	antibiotic resistance protei
pir2:A72527	+	7.00	97.37	700.13	354	hypothetical protein APE2190 - R	pir2:T30222	+	7.00	96.53	692.63	399	sensory protein kinase - Str
pir2:F71721	+	7.00	97.37	700.13	354	hypothetical protein R1120 - R	pir2:B64733	+	7.00	96.51	692.48	400	protein transport protein ho
pir2:T49806	+	7.00	97.37	700.13	354	hypothetical protein B11B22.80	pir2:F90642	+	7.00	96.51	692.48	400	probable integral membrane p
pir2:H96928	+	7.00	97.37	700.13	354	ABC transporter, permease comp	pir2:F83493	+	7.00	96.51	692.48	400	probable integral membrane p
pir2:T15471	-	7.00	97.35	699.95	355	hypothetical protein C09B9.1 -	pir2:A84581	+	7.00	96.47	692.17	402	probable disease resistance
pir2:AE2316	+	7.00	97.33	699.77	356	hypothetical protein AGR_L_636	pir2:T40473	+	7.00	96.45	692.01	403	hypothetical protein SPBC4B4
pir2:D90657	+	7.00	97.26	699.07	360	hypothetical protein ECS0228 [i	pir2:T27825	+	7.00	96.45	692.01	403	hypothetical protein ZK287.7
pir2:D85508	+	7.00	97.26	699.07	360	hypothetical protein Z0259 [imp	pir2:T47621	+	7.00	96.45	692.01	403	bZIP transcription factor-11
pir2:B86347	+	7.00	97.26	699.07	360	hypothetical protein F2438.5 [i	pir2:A83344	-	7.00	96.45	692.01	403	hypothetical protein PA2403
pir2:T29278	-	7.00	97.22	698.72	362	hypothetical protein C34D4.10 -	pir2:E83408	-	7.00	96.45	692.01	403	probable MFS transporter PA1
pir2:H64378	+	7.00	97.20	698.55	363	hypothetical protein M30632 - M	pir2:T51828	+	7.00	96.45	692.01	403	probable photosystem II stab
pir2:ADCBH	+	7.00	97.18	698.38	364	fructose-bisphosphate aldolase	pir2:T34582	+	7.00	96.44	691.86	404	phormone shutdown protein (
pir2:H87184	+	7.00	97.18	698.38	364	5'-phosphoribosyl-5'-aminoimidaz	pir2:B45729	+	7.00	96.44	691.86	404	sulfolipid biosynthesis prot
pir2:E86456	+	7.00	97.18	698.38	364	unknown protein [imported] - Ar	pir2:B46189	+	7.00	96.42	691.86	404	orf within vasotocin gene (T
pir2:AH2429	+	7.00	97.14	698.03	366	hypothetical protein alr4992 [i	pir2:T31912	-	7.00	96.42	691.70	405	hypothetical protein T05H4.2
pir2:H96569	+	7.00	97.14	698.03	366	unknown protein, 54928-56750 [i	pir2:S75974	-	7.00	96.42	691.70	405	hypothetical protein - Synec
pir2:T82370	+	7.00	97.12	697.86	367	heat-inducible transcription re	pir2:A75344	-	7.00	96.40	691.55	406	pi1in biogenesis protein - D
pir2:AD3421	+	7.00	97.10	697.69	368	NADH dehydrogenase (ubiquinone)	pir2:T40582	+	7.00	96.38	691.40	407	hypothetical protein SPBC646
pir2:B64142	+	7.00	97.10	697.69	368	hypothetical protein HI0093 - H	pir2:T34467	+	7.00	96.37	691.24	408	hypothetical protein ZK770.3
pir2:T47445	+	7.00	97.10	697.69	368	hypothetical protein T1B822.100	pir2:A94429	+	7.00	96.35	691.09	409	dihydrolipoamide dehydrogena
pir2:S13721	+	7.00	97.08	697.52	369	Wnt-1 protein precursor, secret	pir2:A95947	-	7.00	96.33	690.94	410	hypothetical protein opuA [
pir2:TVHUT1	+	7.00	97.06	697.35	370	transforming protein int-1 - hu	pir2:T33771	-	7.00	96.33	690.94	410	probable transferase - Srep
pir2:TVMTU1	+	7.00	97.06	697.35	370	transforming protein int-1 - mc	pir2:A71481	+	7.00	96.33	690.94	410	probable poly A polymerase -
pir2:TVMTU1	+	7.00	97.06	697.35	370	transforming protein int-1 - mc	pir2:JX0267	+	7.00	96.28	690.49	413	alpha-1-antitrypsinase S-1 p
pir2:C88474	+	7.00	97.04	697.18	371	transforming protein int-1 pred	pir2:A54968	+	7.00	96.28	690.49	413	alpha-1-antitrypsin precursor
pir2:D86695	+	7.00	97.04	697.18	371	protein C05D10.4 [imported] - C	pir2:S54981	-	7.00	96.28	690.49	413	alpha-1-antitrypsinase Isofo
pir2:E85170	+	7.00	97.02	697.01	372	acetylornithine deacetylase (EC	pir2:TX0154	-	7.00	96.28	690.49	413	alpha-1-antitrypsinase F - r
pir2:AC0253	+	7.00	97.02	697.01	372	hypothetical protein AT4g15450	pir2:D72260	-	7.00	96.28	690.49	413	aminotransferase, class V -
pir2:E89371	-	7.00	96.98	696.84	373	ribonuclease III (EC 3.1.26.3)	pir2:T474714	-	7.00	96.26	690.34	414	hypothetical protein MLCB124
pir2:EB9971	-	7.00	96.98	696.68	374	conserved hypothetical protein	pir2:H72558	+	7.00	96.25	690.19	415	hypothetical protein APE1754
pir2:AC3618	-	7.00	96.97	696.51	375	alcohol dehydrogenase (EC 1.1.1.1	pir2:T13435	+	7.00	96.25	690.19	415	hypothetical protein T17A13.
pir2:C90444	+	7.00	96.95	696.34	376	hypothetical protein acac-8 [im	pir2:B86919	+	7.00	96.21	689.89	417	probable seryl-tRNA synthase
pir2:D81980	+	7.00	96.93	696.18	377	carbamoyl-phosphate synthase (g	pir2:T34459	+	7.00	96.21	689.89	417	hypothetical protein T19H12
pir2:F81034	+	7.00	96.93	696.18	377	carbamoyl-phosphate synthase, S	pir2:F95953	+	7.00	96.21	689.89	417	probable glycosyltransferase
pir2:B38178	-	7.00	96.91	696.01	378	telA protein - plasmid RK2	pir2:E91037	+	7.00	96.20	689.74	418	hypothetical protein ECS3269
pir2:FWB7W	+	7.00	96.89	695.84	379	tryptophan-tRNA ligase (EC 6.1.1	pir2:B65013	+	7.00	96.20	689.74	418	hypothetical protein Z3389 -
pir2:S17918	+	7.00	96.89	695.84	379	H+-transporting ATP synthase (E	pir2:B65013	+	7.00	96.20	689.74	418	hypothetical protein b2655 [
pir2:F76029	-	7.00	96.89	695.84	379	hypothetical protein - Synecoc	pir2:T30652	+	7.00	96.18	689.59	419	probable serS protein - Myco
pir2:JC2579	+	7.00	96.87	695.68	380	manganese peroxidase (EC 1.11.1	pir2:A90888	+	7.00	96.18	689.59	419	hypothetical protein ECS2073
pir2:T46884	+	7.00	96.85	695.51	381	3-methyl-2-oxobutanoate dehydro	pir2:H86230	+	7.00	96.18	689.59	419	hypothetical protein [import
pir2:AJ9992	+	7.00	96.83	695.35	382	hypothetical protein orf382 [im	pir2:T42929	+	7.00	96.18	689.59	419	hypothetical protein T16124
pir2:JC5753	+	7.00	96.82	695.19	383	pepsinogen C - African clawed f	pir2:H85729	+	7.00	96.18	689.59	419	hypothetical protein Z2242 [
pir2:T38443	+	7.00	96.82	695.19	383	hypothetical protein SPAC27D7.0	pir2:DC8CD	+	7.00	96.16	689.44	420	diaminopimelate decarboxylas
pir2:T38443	+	7.00	96.82	695.19	383	hypothetical protein [imported]	pir2:G91090	+	7.00	96.16	689.44	420	diaminopimelate decarboxylas
pir2:H86232	+	7.00	96.82	695.19	383	hypothetical protein [imported]	pir2:G91090	+	7.00	96.16	689.44	420	diaminopimelate decarboxylas
pir2:AB6182	+	7.00	96.82	695.19	383	hypothetical protein [imported]	pir2:DE8936	+	7.00	96.16	689.44	420	diaminopimelate decarboxylas
pir2:E75367	-	7.00	96.80	695.02	384	probable oxidoreductase - Deinc	pir2:DE8936	-	7.00	96.14	689.30	421	acyl-CoA dehydrogenase (EC 1
pir2:A39314	+	7.00	96.80	695.02	384	gastricsin (EC 3.4.23.3) precu	pir2:AS5724	+	7.00	96.14	689.30	421	acyl-CoA dehydrogenase (EC 1
pir2:JC5206	+	7.00	96.80	695.02	384	hypothetical 43.0K protein - Ch	pir2:C72416	+	7.00	96.14	689.30	421	UDP-N-acetylglucosamine 1-ca
pir2:F90649	-	7.00	96.78	694.86	385	hypothetical protein ECS0166 [i	pir2:AD2496	+	7.00	96.14	689.30	421	transposase all17148 [importe
pir2:AD0528	-	7.00	96.78	694.86	385	conserved hypothetical protein	pir2:S75970	-	7.00	96.14	689.30	421	hypothetical protein - Synec
pir2:E96669	+	7.00	96.78	694.86	385	protein FLN9.19 [imported] - A	pir2:T13915	+	7.00	96.14	689.30	421	hypothetical protein C15H11.
pir2:T29315	+	7.00	96.78	694.86	385	hypothetical protein F36D4.5 -	pir2:C70815	+	7.00	96.11	689.00	423	probable hiss protein - Myco
pir2:G71467	-	7.00	96.76	694.70	386	probable succinyl-coa synthetas	pir2:T08412	-	7.00	96.09	688.86	424	hypothetical protein F18B3.1

pir2:D84677	+	7.00	96.09	688.86	424	!	hypothetical protein At2g27820	pir2:T51030	-	7.00	95.50	683.69	461	!	probable acetylornithine ami
pir1:I50898	-	7.00	96.08	688.71	425	!	inhibin beta-A chain precursor	pir2:B69676	+	7.00	95.50	683.69	461	!	alkaline phosphatase (EC 3.1
pir2:I47072	-	7.00	96.08	688.71	425	!	inhibin beta-A chain precursor	pir2:A40552	+	7.00	95.49	683.56	462	!	bindin fertilization specifi
pir2:S73739	+	7.00	96.04	688.42	427	!	histidine--trNA ligase (EC 6.1	pir2:AB3893	-	7.00	95.49	683.56	462	!	two-component sensor histidi
pir2:T14421	+	7.00	96.04	688.42	427	!	histidyl-trNA synthase [import	pir1:WMECHP	+	7.00	95.47	683.43	463	!	hexose phosphate transport p
pir2:I875213	-	7.00	96.04	688.42	427	!	s-locus-specific glycoprotein -	pir2:T10065	+	7.00	95.47	683.43	463	!	phosphoadenylyl-sulfate redu
pir2:D83347	+	7.00	96.04	688.42	427	!	pmba protein (tlde protein) PAE	pir2:C91204	+	7.00	95.47	683.43	463	!	hexose phosphate transport p
pir2:JCA4986	+	7.00	96.01	688.13	429	!	site-specific DNA-methyltransf	pir2:E86050	+	7.00	95.47	683.43	463	!	hexose phosphate transport p
pir2:J75013	+	7.00	96.01	688.13	429	!	phosphoribosylglycinamide formy	pir2:S44878	+	7.00	95.42	683.03	466	!	2C262.6 protein - Caenorhabd
pir2:T05011	+	7.00	96.01	688.13	429	!	hypothetical protein T19P19.130	pir2:G87085	-	7.00	95.36	682.51	470	!	arginosuccinate lyase [impor
pir2:T28400	+	7.00	96.01	688.13	429	!	ORF MSV239 leucine rich repeat	pir2:F70621	-	7.00	95.36	682.51	470	!	probable argH protein - Myco
pir2:T04668	+	7.00	95.99	687.99	430	!	phosphoserine transaminase homc	pir2:F82302	+	7.00	95.36	682.51	470	!	probable phosphoglucomutase/
pir2:T43437	+	7.00	95.99	687.99	430	!	hypothetical protein MLCB373.07	pir2:T87050	+	7.00	95.33	682.24	472	!	diaminopimelate decarboxylas
pir2:T567161	-	7.00	95.99	687.99	430	!	hypothetical protein YOR264w -	pir2:C70779	+	7.00	95.32	682.11	473	!	probable accD6 protein - Ydc
pir2:T46099	+	7.00	95.98	687.99	431	!	hypothetical protein T25B15.60	pir2:S51256	-	7.00	95.32	682.11	473	!	probable membrane protein YD
pir1:WJHU2G	+	7.00	95.98	687.84	431	!	homeotic protein Hox B3 - human	pir2:A86442	+	7.00	95.32	682.11	473	!	hypothetical protein F5W6.26
pir2:T20263	+	7.00	95.98	687.84	431	!	hypothetical protein C56A3.6 -	pir2:T38905	-	7.00	95.30	681.99	474	!	probable valine--pyruvate tr
pir2:T71138	+	7.00	95.95	687.56	433	!	phosphoribosylglycinamide formy	pir2:T45818	+	7.00	95.30	681.99	474	!	hypothetical protein F2809.2
pir2:S20963	+	7.00	95.95	687.56	433	!	homeotic protein Hox B3 - mouse	pir2:T50258	+	7.00	95.30	681.99	474	!	hypothetical coiled-coil pro
pir2:T72241	+	7.00	95.86	686.85	438	!	fixC protein - Thermotoga marit	pir2:H71660	+	7.00	95.29	681.86	475	!	nitrogen assimilation regula
pir2:A83544	+	7.00	95.86	686.85	438	!	probable transporter PA0809 [lin	pir2:T01352	+	7.00	95.29	681.86	475	!	hypothetical protein F6N15.2
pir2:S67921	-	7.00	95.85	686.71	439	!	multiple ligand-binding protein	pir2:T41935	+	7.00	95.26	681.60	477	!	capsid protein - human herpes
pir2:A59227	-	7.00	95.85	686.71	439	!	site-specific recombinase gcr -	pir2:T45722	+	7.00	95.26	681.60	477	!	hypothetical protein F1P2.17
pir2:T43798	+	7.00	95.85	686.71	439	!	hypothetical protein F2809.50 -	pir2:C83739	-	7.00	95.24	681.47	478	!	hypothetical protein BH0715
pir2:F70337	+	7.00	95.85	686.71	439	!	hypothetical protein aq.414 - A	pir2:G86207	+	7.00	95.23	681.34	479	!	hypothetical protein [import
pir2:T64090	+	7.00	95.83	686.57	440	!	dicarboxylate transport protein	pir2:T00971	+	7.00	95.21	681.21	480	!	probable disease resistance
pir2:H66667	+	7.00	95.83	686.57	440	!	UDP-N-acetylmuramoylalanyl-D-gl	pir2:A89102	+	7.00	95.20	681.09	481	!	protein F25E5.4 [imported] -
pir2:T71106	+	7.00	95.83	686.57	440	!	hypothetical protein PH0625 - F	pir2:A44268	-	7.00	95.18	680.96	482	!	cactus - fruit fly (Drosophi
pir2:A83211	+	7.00	95.80	686.29	442	!	conserved hypothetical protein	pir2:S36184	+	7.00	95.14	680.58	485	!	translation elongation facto
pir2:GF1866	+	7.00	95.78	686.15	443	!	hypothetical protein px01.95 -	pir2:B55886	+	7.00	95.13	680.45	486	!	dopamine receptor D1B - chic
pir2:A59102	+	7.00	95.78	686.15	443	!	hypothetical protein all0479 [li	pir2:T21481	+	7.00	95.13	680.45	486	!	hypothetical protein F28C6.1
pir2:A75319	-	7.00	95.77	686.01	444	!	phosphoglucomutase and phospho	pir2:AC1726	+	7.00	95.13	680.45	486	!	amino acid ABC transporter,
pir2:B83759	-	7.00	95.77	686.01	444	!	alkaline phosphatase BH0874 [lin	pir2:T16160	+	7.00	95.10	680.20	488	!	hypothetical protein F4E83.2
pir2:T23725	-	7.00	95.75	685.87	445	!	hypothetical protein M0585.1 -	pir2:C86491	+	7.00	95.04	679.70	492	!	Glu tRNA Gln amidotransferr
pir2:T34352	+	7.00	95.75	685.87	445	!	hypothetical protein T12A2.1 -	pir2:H72130	+	7.00	95.04	679.70	492	!	glutamyl-tRNA(Gln) amidotran
pir2:T28015	+	7.00	95.75	685.87	445	!	hypothetical protein ZK822.4 -	pir2:AB2434	+	7.00	95.02	679.58	493	!	hypothetical protein all5026
pir2:S01187	-	7.00	95.74	685.73	446	!	NADH dehydrogenase (ubiquinone)	pir2:E71055	+	7.00	95.01	679.46	494	!	hypothetical protein PH1139
pir2:T25797	-	7.00	95.74	685.73	446	!	NADH dehydrogenase (ubiquinone)	pir2:G96938	-	7.00	94.95	678.96	498	!	sensory transduction histidi
pir2:F87337	+	7.00	95.72	685.59	447	!	diaminopimelate decarboxylase (	pir1:S32331	-	7.00	94.92	678.72	500	!	proline/betaine transport pr
pir1:A70773	+	7.00	95.72	685.59	447	!	2-nitrotoluene dioxygenase (EC	pir2:E91265	-	7.00	94.92	678.72	500	!	proline/betaine transport pr
pir2:JC5352	+	7.00	95.72	685.59	447	!	hypothetical protein SPC13Cl.1	pir2:B6106	-	7.00	94.92	678.72	500	!	proline/betaine transport pr
pir2:T39414	+	7.00	95.70	685.45	448	!	conserved hypothetical protein	pir2:AB1022	-	7.00	94.92	678.72	500	!	Prop [imported] - Salmonella
pir2:D81740	+	7.00	95.69	685.32	449	!	competence factor TransComB [lin	pir2:A4269	-	7.00	94.92	678.72	500	!	ankyrin repeat acidic protei
pir2:D95004	+	7.00	95.69	685.32	449	!	transport protein ComB comb [lin	pir2:A4268	-	7.00	94.92	678.72	500	!	cactus - fruit fly (Drosophi
pir2:D97877	+	7.00	95.69	685.32	449	!	hypothetical protein BH3535 [lin	pir2:T02134	+	7.00	94.91	678.60	501	!	hypothetical protein F8K4.9
pir2:G84091	+	7.00	95.66	685.04	451	!	gamma-aminobutyric acid/benzodi	pir2:A70988	+	7.00	94.90	678.47	502	!	hypothetical protein Rv1760
pir1:AC8062	-	7.00	95.66	685.04	451	!	gamma-aminobutyric acid/benzodi	pir2:T00483	-	7.00	94.90	678.47	502	!	hypothetical protein At2g350
pir2:JC0370	-	7.00	95.66	685.04	451	!	probable glutamine synthetase (	pir2:T05135	+	7.00	94.90	678.47	502	!	hypothetical protein F7H19.2
pir2:T157947	-	7.00	95.66	685.04	451	!	probable glutamine synthetase (	pir2:E90325	+	7.00	94.88	678.35	503	!	benzoylformate decarboxylase
pir2:C97382	-	7.00	95.64	684.91	452	!	nicotinic acetylcholine recepto	pir2:T29868	+	7.00	94.88	678.35	503	!	hypothetical protein F1062.8
pir2:A35721	+	7.00	95.64	684.91	452	!	conserved hypothetical protein	pir2:S33194	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir2:C69140	+	7.00	95.64	684.91	452	!	carbamoyl-phosphate synthase (E	pir2:S33186	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir2:A43224	+	7.00	95.62	684.77	453	!	conserved hypothetical protein	pir2:S33190	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir2:AHO126	+	7.00	95.61	684.63	454	!	probable ubiquitin activating e	pir2:S33191	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir2:T52523	+	7.00	95.59	684.50	455	!	gamma-aminobutyric acid/benzodi	pir2:S33189	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir1:CHCRA1	-	7.00	95.59	684.50	455	!	gamma-aminobutyric acid/benzodi	pir2:S33188	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir2:JQ0158	-	7.00	95.59	684.50	455	!	gamma-aminobutyric acid/benzodi	pir2:S33187	+	7.00	94.87	678.23	504	!	phase-1 flagellin - Salmonel
pir2:A30662	-	7.00	95.59	684.50	455	!	gamma-aminobutyric acid/benzodi	pir2:S33192	+	7.00	94.83	677.87	507	!	phase-1 flagellin - Salmonel
pir2:A70717	+	7.00	95.59	684.50	455	!	hypothetical protein YDR393w -	pir2:S33185	+	7.00	94.83	677.87	507	!	phase-1 flagellin - Salmonel
pir2:T32189	+	7.00	95.59	684.50	455	!	zinc finger protein nhr-55 - Ca	pir2:T02093	+	7.00	94.83	677.87	507	!	hypothetical protein D1046.1
pir2:A60652	-	7.00	95.58	684.36	456	!	gamma-aminobutyric acid/benzodi	pir2:A53465	+	7.00	94.81	677.75	508	!	phase 1 flagellin - Salmonel
pir2:A27142	-	7.00	95.58	684.36	456	!	hypothetical protein 456 (tufa	pir2:C82138	+	7.00	94.81	677.75	508	!	conserved hypothetical prote
pir2:S69677	+	7.00	95.58	684.36	456	!	hypothetical protein YDR393w -	pir2:D69832	+	7.00	94.80	677.63	509	!	probable Rieske [2Fe-2S] iro
pir2:T22173	+	7.00	95.56	684.23	457	!	hypothetical protein F4E5.3 -	pir2:T39542	+	7.00	94.80	677.63	509	!	hypothetical protein SPC16A
pir2:T28334	+	7.00	95.56	684.23	457	!	ORF MSV173 probable serine/thre	pir2:JC5880	+	7.00	94.77	677.39	511	!	L-2,4-diaminobutyrate decarb
pir2:T17178	+	7.00	95.53	683.96	459	!	NADH dehydrogenase (ubiquinone)	pir1:OAHU6	+	7.00	94.76	677.27	512	!	aryl hydrocarbon (benzo[a]lpy
pir2:T17181	+	7.00	95.53	683.96	459	!	NADH dehydrogenase (ubiquinone)	pir2:S21761	+	7.00	94.76	677.27	512	!	aryl hydrocarbon (benzo[a]lpy
pir2:T17175	+	7.00	95.53	683.96	459	!	NADH dehydrogenase (ubiquinone)								
pir2:T39717	+	7.00	95.53	683.96	459	!	probable dna primase large subu								
pir2:T27657	+	7.00	95.53	683.96	459	!	hypothetical protein ZK1037.1 -								
pir2:AB2600	-	7.00	95.52	683.83	460	!	glutamine synthetase glna [impd								

pir2:H84707	7.00	94.76	677.27	512	probable ferrocyclase precursor	pir2:T11628	7.00	93.49	666.49	612	MVP1 protein - fission yeast
pir2:G86459	7.00	94.76	677.27	512	Hypothetical 55.6 kDa protein -	pir1:S74461	7.00	93.48	666.39	613	ABC transporter sir1494 - Sy
pir1:I58311	7.00	94.74	677.15	513	HMG-box containing protein 1 -	pir2:T15489	7.00	93.48	666.39	613	hypothetical protein C14F11
pir2:E96533	7.00	94.74	677.15	513	probable fucosyltransferase [im	pir2:E73094	7.00	93.47	666.30	614	prolyl endopeptidase PAB0762
pir2:B96925	7.00	94.71	676.92	515	hypothetical protein T8K14.17 [i	pir2:F96791	7.00	93.47	666.30	614	hypothetical protein F15M4.1
pir2:T05396	7.00	94.67	676.56	518	hypothetical protein F4K10.90 -	pir2:A32608	7.00	93.47	666.30	614	thyroid hormone receptor-rel
pir2:AD0663	7.00	94.66	676.41	519	hypothetical protein STY1413 [i	pir2:A3188	7.00	93.46	666.20	615	hypothetical protein PA3670
pir2:H84202	7.00	94.63	676.21	521	Na+/H+ antiporter [imported] -	pir2:H96732	7.00	93.46	666.20	615	hypothetical protein F15H11.
pir2:F17302	7.00	94.61	675.98	523	asparagine--tRNA ligase (EC 6.1	pir2:B71071	7.00	93.43	666.00	617	probable prolyl endopeptidase
pir2:D64555	7.00	94.61	675.98	523	conserved hypothetical integral	pir2:S52797	7.00	93.43	666.00	617	mufl protein - human
pir2:A53467	7.00	94.55	675.51	527	protein kinase SNF1 homolog wpx	pir2:S08446	7.00	93.42	665.91	618	citrolysin-related protein 1
pir2:A22605	7.00	94.51	675.17	530	halolysin (EC 3.4.21.-) - Natri	pir2:T28446	7.00	93.41	665.81	619	hypothetical protein ZC15.2
pir2:S22340	7.00	94.51	675.17	530	seeligeriolysin - Listeria seel	pir2:B83878	7.00	93.36	665.42	623	acetoaldehyde dehydrogenase operon
pir2:T04463	7.00	94.50	675.06	531	hypothetical protein F4D11.170	pir2:B83878	7.00	93.36	665.42	623	hypothetical protein Cj004c
pir2:D96710	7.00	94.47	674.83	533	F24J5.16 [imported] - Arabidops	pir2:G81420	7.00	93.36	665.42	623	hypothetical protein YDL203c
pir2:T32925	7.00	94.47	674.83	533	hypothetical protein SPAC1367.1	pir2:S67762	7.00	93.36	665.42	623	hypothetical protein C41C4.3
pir1:A43162	7.00	94.46	674.71	534	3',5'-cyclic-nucleotide phospho	pir2:T19876	7.00	93.36	665.42	623	hypothetical protein C41C4.3
pir2:T39903	7.00	94.46	674.71	534	serine-rich protein - fission y	pir2:T28423	7.00	93.35	665.33	624	ORF MSV261 leucine rich repe
pir2:B33485	7.00	94.42	674.37	537	spore coat protein SP70 - slime	pir2:T03837	7.00	93.34	665.23	625	surface lipoprotein - Mycopl
pir2:T08640	7.00	94.40	674.26	538	hypothetical protein b2628 - Es	pir2:B70749	7.00	93.33	665.14	626	probable Acyl-CoA Synthetase
pir2:T12704	7.00	94.38	674.04	540	leucine-rich protein - common s	pir2:T03821	7.00	93.32	665.04	627	prolipoprotein p65 - Mycopla
pir2:T08777	7.00	94.35	673.81	542	pectate lyase (EC 4.2.2.2) F28	pir2:T31798	7.00	93.28	664.76	630	hypothetical protein R02F11.
pir2:T06728	7.00	94.35	673.81	542	scarecrow-like 7 (SCU7) - Arabi	pir2:T31017	7.00	93.28	664.76	630	hypothetical protein 630A -
pir2:T46142	7.00	94.35	673.81	542	hypothetical protein F19C14.9 [i	pir2:A83447	7.00	93.19	664.00	634	heat shock protein Hsp90 PA15
pir2:E96616	7.00	94.34	673.70	543	probable ATP-dependent RNA heli	pir2:T31383	7.00	93.19	664.00	638	receptor protein kinase-like
pir2:E87010	7.00	94.33	673.59	544	hypothetical protein T12P18.13	pir2:C84726	7.00	93.16	663.72	641	probable receptor-like prote
pir2:T02578	7.00	94.31	673.48	545	hypothetical protein At2g39240	pir2:A83268	7.00	93.15	663.63	642	probable soluble lytic trans
pir2:F96663	7.00	94.27	673.15	548	hypothetical protein T12P18.13	pir2:C95777	7.00	93.15	663.63	642	F25M4.24 [imported] - Arabid
pir1:S39533	7.00	94.26	673.03	549	phosphoprotein phosphatase (EC	pir2:S55610	7.00	93.14	663.54	643	polyprotein - equine herpesv
pir2:T41744	7.00	94.26	673.03	549	hypothetical protein F15U1.40 -	pir2:T13225	7.00	93.14	663.54	643	No autoantigen 60K homolog -
pir2:B98302	7.00	94.26	673.03	549	periplasmic dipeptide transport	pir2:F97787	7.00	93.14	663.54	643	sodium/pantothenate symporte
pir2:A22981	7.00	94.26	673.03	549	hypothetical protein At3455 [i	pir2:A86623	7.00	93.13	663.44	644	transketolase [imported] - C
pir2:S64314	7.00	94.24	672.81	551	probable membrane protein VGR02	pir2:A72002	7.00	93.13	663.44	644	1-deoxyxylulose-5-phosphate
pir2:D95139	7.00	94.18	672.38	555	DNA repair and genetic recombin	pir2:T11137	7.00	93.12	663.35	645	NADH dehydrogenase (ubiquino
pir2:C98007	7.00	94.18	672.38	555	DNA repair and genetic recombin	pir2:T28867	7.00	93.12	663.35	645	hypothetical protein R03H4.5
pir2:T07116	7.00	94.16	672.16	557	protoporphyrinogen oxidase (EC	pir2:T33141	7.00	93.10	663.17	647	hypothetical protein SPAC8C9
pir1:G64246	7.00	94.11	671.73	561	conserved hypothetical protein	pir2:S08450	7.00	93.10	663.17	647	steroid hormone receptor hom
pir2:T34313	7.00	94.09	671.62	562	hypothetical protein K03A1.2 -	pir2:S51409	7.00	93.08	663.07	648	hypothetical protein YLR273c
pir2:A71513	7.00	94.08	671.51	563	probable arginyl tRNA transfer	pir2:E98215	7.00	93.07	662.98	649	hypothetical protein AGR_L30
pir2:F81670	7.00	94.08	671.51	563	arginyl-tRNA synthetase TC0739	pir2:T10219	7.00	93.05	662.80	651	protein kinase homolog T30C3
pir1:HM1676	7.00	94.04	671.19	566	hemagglutinin precursor - Infl	pir2:T10219	7.00	93.05	662.80	651	protein kinase homolog T30C3
pir2:T31964	7.00	94.04	671.19	566	hypothetical protein C33C12.3 -	pir2:B84690	7.00	93.02	662.52	654	hypothetical protein At2g288
pir2:S74633	7.00	94.03	671.08	567	high affinity sulfate transport	pir2:T30044	7.00	93.01	662.43	655	hypothetical protein C16H3.3
pir2:AG2008	7.00	94.03	671.08	567	hypothetical protein alr1621 [i	pir2:AD3071	7.00	93.01	662.43	655	hypothetical protein At4A192
pir2:S51275	7.00	94.02	670.98	568	DNA polymerase - phage CP-1	pir2:T37941	7.00	93.00	662.34	656	conserved hypothetical prote
pir2:AI2474	7.00	93.99	670.77	570	hypothetical protein alr5353 [i	pir2:S43415	7.00	92.99	662.25	657	histidine ammonia-lyase (EC
pir2:E90183	7.00	93.97	670.55	572	hypothetical protein S800395 [i	pir2:T33568	7.00	92.98	662.16	658	hypothetical protein R160.7
pir2:C86779	7.00	93.97	670.55	572	conserved hypothetical protein	pir2:JC4910	7.00	92.97	662.07	659	heparitin-sulfate lyase (EC
pir2:AC3651	7.00	93.97	670.55	572	succinoglycan biosynthesis tran	pir2:T11641	7.00	92.97	662.07	659	hypothetical protein SPAC3G9
pir2:C86806	7.00	93.96	670.45	573	hypothetical protein busAB [imp	pir2:T40383	7.00	92.97	662.07	659	hypothetical protein SPBC3E7
pir2:T51239	7.00	93.96	670.45	573	scarecrow-like protein 8 [import	pir2:B70662	7.00	92.95	661.98	660	probable membraneprotein - M
pir2:D70114	7.00	93.94	670.34	574	PTS system, maltose and glucose	pir2:G86428	7.00	92.93	661.80	662	F26G16.2 protein - Arabidops
pir2:AC0109	7.00	93.91	670.03	577	single-stranded-DNA-specific ex	pir2:G86210	7.00	92.93	661.80	662	hypothetical protein [import
pir2:A83771	7.00	93.91	670.03	577	SNF2 helicase BH0369 [imported]	pir2:T04856	7.00	92.93	661.80	662	hypothetical protein [import
pir2:T37248	7.00	93.88	669.82	579	probable matrix metalloprotein	pir2:AF3418	7.00	92.92	661.71	663	cytochrome c-type biogenesis
pir2:F84828	7.00	93.87	669.72	580	probable laccase (diphenol oxid	pir2:S44756	7.00	92.91	661.62	664	probable protein disulfide-1
pir2:T09157	7.00	93.83	669.41	583	phosphoglucutamate precursor, c	pir2:AG1177	7.00	92.90	661.53	665	transketolase homolog lmo034
pir2:C89785	7.00	93.79	669.00	587	hypothetical protein SA0220 [im	pir2:A11477	7.00	92.89	661.44	666	transketolase homolog lln036
pir2:A49618	7.00	93.76	668.89	588	probable ataxia-telangiectasia	pir2:E71565	7.00	92.89	661.44	666	probable glycogen hydrolase
pir2:T18239	7.00	93.75	668.79	589	transcription effector - yeast	pir2:G81717	7.00	92.89	661.44	666	glycosyl hydrolase family pr
pir2:S63193	7.00	93.75	668.69	590	hypothetical protein YNL227c -	pir2:T11778	7.00	92.87	661.26	668	phosphoglycerate transport r
pir2:T27710	7.00	93.71	668.38	593	hypothetical protein ZK1193.5 -	pir2:AH0806	7.00	92.87	661.26	668	phosphoglycerate transport r
pir2:T41007	7.00	93.69	668.18	595	atp dependent helicase - fission	pir2:S64795	7.00	92.86	661.17	669	suppressor protein Fps1 - ye
pir2:T47673	7.00	93.67	667.98	597	hypothetical protein T26112.220	pir2:JC5662	7.00	92.85	661.17	669	hepatonate-derived growth fact
pir2:T02795	7.00	93.65	667.88	598	probable membrane protein L549	pir1:A30882	7.00	92.85	661.09	670	arachidonate 5-lipoxygenase
pir2:T32166	7.00	93.65	667.88	598	hypothetical protein C31A8.8 -	pir2:D86324	7.00	92.84	661.00	671	protein F14D16.25 [imported]
pir2:T25207	7.00	93.65	667.88	598	hypothetical protein T23G7.1 -	pir2:H81976	7.00	92.84	661.00	671	probable prolyl oligopeptida
pir2:T25207	7.00	93.65	667.88	598	hypothetical protein T23G7.1 -	pir2:C81035	7.00	92.84	661.00	671	prolyl oligopeptidase family
pir1:A64235	7.00	93.64	667.78	599	cytadherence-accessory protein	pir2:T23013	7.00	92.84	661.00	671	hypothetical protein F59F5.3
pir2:S75048	7.00	93.64	667.78	599	aspartate--tRNA ligase (EC 6.1	pir2:AF1143	7.00	92.82	660.82	673	internalin protein homolog 1
pir2:T34391	7.00	93.54	666.88	608	hypothetical protein T26N5.1 -	pir2:IA9479	7.00	92.81	660.73	674	arachidonate 5-lipoxygenase
pir2:AI3325	7.00	93.54	666.88	608	hypothetical protein T26N5.1 -	pir2:C81505	7.00	92.81	660.73	674	hypothetical protein CPI075
pir2:T18271	7.00	93.50	666.58	611	hypothetical protein ORF10 - Me	pir2:T4261	7.00	92.81	660.73	674	hypothetical protein T1E22.1
						pir2:H86198	7.00	92.79	660.64	675	hypothetical protein [import

pir2:T40680	7.00	92.79	660.64	675	hypothetical protein SPAC776.1d	pir2:D90099	7.00	91.83	652.64	773	hypothetical protein orf773
pir2:A45515	7.00	92.78	660.56	676	dnaK-type molecular chaperone	pir2:C96805	7.00	91.78	652.18	779	hypothetical protein T5M16.1
pir2:S73798	7.00	92.77	660.47	677	MG260 homolog H91_orf677 - MycQ	pir2:T48189	7.00	91.77	652.11	780	probable transporter protein
pir2:H86208	7.00	92.77	660.47	677	protein F22G5.26 [imported] - A	pir2:S58235	7.00	91.69	651.44	789	endo-1,4-beta-xylanase (EC 3
pir2:E72033	7.00	92.74	660.20	680	hypothetical protein - Chlamydo	pir1:WMBEA2	7.00	91.68	651.36	790	ribonucleoside-diphosphate r
pir2:H86590	7.00	92.74	660.20	680	hypothetical protein CPJ0796 [i	pir2:T41573	7.00	91.67	651.29	791	hypothetical protein SPCC736
pir2:S90095	7.00	92.70	659.86	684	DNA gyrase subunit B [imported]	pir2:T38490	7.00	91.61	650.77	798	hypothetical protein SPAC298
pir2:T06700	7.00	92.68	659.68	686	hypothetical protein T29H11.60	pir2:JH0797	7.00	91.60	650.70	799	castor protein - fruit fly (C
pir2:S74621	7.00	92.67	659.60	687	carbon dioxide concentrating me	pir1:TVHU2F	7.00	91.59	650.62	800	fibroblast growth factor rec
pir2:A56360	7.00	92.66	659.51	688	zinc finger protein (Clone 18)	pir2:A48991	7.00	91.59	650.62	800	heparin-binding growth facto
pir2:AE1132	7.00	92.65	659.42	689	probable membrane associated li	pir2:I53363	7.00	91.58	650.55	801	fibroblast growth factor rec
pir2:DV5487	7.00	92.64	659.34	690	v-type ATP synthase, I subunit	pir2:AE1960	7.00	91.58	650.55	801	fibroblast growth factor rec
pir2:S64904	7.00	92.61	659.08	693	probable membrane protein YLR07	pir1:TVHUF3	7.00	91.54	650.19	806	fibroblast growth factor rec
pir2:T42659	7.00	92.58	658.83	697	hypothetical protein DKF2p434K1	pir2:T14248	7.00	91.54	650.19	806	sucrose synthase [imported]
pir2:T34704	7.00	92.57	658.74	697	probable transferase - Streptoc	pir2:T51232	7.00	91.52	650.04	808	scarcrow-like protein 14 [i
pir2:T21148	7.00	92.52	658.32	702	hypothetical protein F20D1.7 -	pir2:AY1419	7.00	91.47	649.61	814	probable allene oxide syntha
pir2:A65383	7.00	92.52	658.15	704	76.4K protein kinase homolog F4	pir2:F96586	7.00	91.43	649.32	818	hypothetical protein F20D21.
pir2:T50565	7.00	92.50	658.15	704	hypothetical protein F22K18.110	pir2:T34232	7.00	91.41	649.11	821	hypothetical protein F2086.8
pir2:S31612	7.00	92.50	658.15	704	probable membrane protein YGL24	pir2:T34472	7.00	91.39	648.97	823	hypothetical protein W03B1.2
pir2:S38168	7.00	92.48	657.98	706	hypothetical protein YKR090w -	pir2:T10615	7.00	91.38	648.90	824	hypothetical protein F21C20.
pir2:T88418	7.00	92.47	657.90	707	ORF MSV257 leucine rich repeat	pir2:S26706	7.00	91.37	648.83	825	transcription factor SWI3 -
pir2:E84950	7.00	92.46	657.81	708	phosphate acetyltransferase (EC	pir2:AC0086	7.00	91.36	648.76	826	outer membrane usher protein
pir2:B84853	7.00	92.39	657.23	715	hypothetical protein At2g42370	pir2:C96639	7.00	91.35	648.61	828	protein TIF9.14 [imported] -
pir2:S65175	7.00	92.39	657.23	715	hypothetical protein YPL164c -	pir2:S50246	7.00	91.34	648.54	829	SIP4 protein - yeast (Saccha
pir2:T95221	7.00	92.34	656.82	720	alpha-galactosidase [imported]	pir2:T32651	7.00	91.33	648.47	830	probable integral membrane p
pir2:G98085	7.00	92.34	656.82	720	alpha-galactosidase (EC 3.2.1.2	pir2:S50810	7.00	91.33	648.47	830	probable membrane protein YJ
pir1:Q0BE15	7.00	92.30	656.49	724	UL89 protein - human cytomegalo	pir2:AS9283	7.00	91.32	648.40	831	unconventional myosin-A - To
pir2:B71404	7.00	92.30	656.49	724	hypothetical protein - Arabidop	pir2:T12827	7.00	91.25	647.85	839	hypothetical protein yono -
pir2:T01268	7.00	92.29	656.41	725	leucine-rich repeat transmembra	pir2:AT87341	7.00	91.23	647.64	842	sensory box/GDPF family pro
pir2:T90435	7.00	92.28	656.33	726	hypothetical protein dppA [impc	pir2:HT72204	7.00	91.22	647.57	843	pullulanase - Thermotoga mar
pir2:T18665	7.00	92.27	656.15	727	hypothetical protein B0035.6 -	pir2:T32651	7.00	91.22	647.57	843	hypothetical protein F39c12.
pir2:B83805	7.00	92.26	656.17	728	GTP pyrophosphokinase (stringe	pir2:C91212	7.00	91.21	647.50	844	probable outer membrane ushe
pir2:D86278	7.00	92.26	656.17	728	hypothetical protein F14L17.16	pir2:C86058	7.00	91.21	647.50	844	probable fibrial usher Z522
pir2:A83536	7.00	92.23	655.92	731	conserved hypothetical protein	pir2:JC4054	7.00	91.21	647.50	844	membrane alanyl aminopeptida
pir1:JCA655	7.00	92.22	655.84	732	acylaminoacyl-peptidase (EC 3.4	pir2:S47274	7.00	91.21	647.50	844	membrane alanyl aminopeptida
pir1:J00132	7.00	92.22	655.84	732	acylaminoacyl-peptidase (EC 3.4	pir2:T07039	7.00	91.20	647.43	845	Hcr9-0 protein - tomato
pir1:S07624	7.00	92.22	655.84	732	acylaminoacyl-peptidase (EC 3.4	pir2:G86708	7.00	91.18	647.22	848	aconitate hydratase (EC 4.2.
pir2:S52210	7.00	92.22	655.84	732	coat protein VP1 - muscovy duck	pir2:S12159	7.00	91.15	647.02	851	env protein - human immunode
pir2:S56277	7.00	92.21	655.76	733	probable membrane protein YFR02	pir2:AD1427	7.00	91.15	647.02	851	internalin, probable peptido
pir2:A83636	7.00	92.21	655.76	733	sulfite reductase (NADPH) flavo	pir2:AY2343	7.00	91.14	646.95	852	hypothetical protein - Therm
pir2:T40510	7.00	92.18	655.52	736	beta transducin - fission yeast	pir2:T41336	7.00	91.12	646.75	855	probable nitrogen regulatory
pir2:S72442	7.00	92.17	655.44	737	actin-fragmin kinase - slime mc	pir2:T00258	7.00	91.09	646.54	858	hypothetical protein KIAA060
pir1:TFH096	7.00	92.16	655.36	738	melanotransferrin precursor - f	pir2:A82044	7.00	91.08	646.41	860	adenylate cyclase N [imported]
pir2:T02996	7.00	92.15	655.28	739	N-ethylmaleimide sensitive fusi	pir2:HB7556	7.00	91.05	646.40	863	aminopeptidase N [imported]
pir2:B84741	7.00	92.14	655.20	740	hypothetical protein At2q33080	pir1:A28443	7.00	91.00	645.80	869	phosphatidylethanolamine N-m
pir2:T05250	7.00	92.13	655.12	741	probable disease resistance pro	pir2:S53098	7.00	91.00	645.80	869	envelope polyprotein - human
pir2:G90124	7.00	92.10	654.89	744	hypothetical protein orf744 [im	pir2:S65158	7.00	90.99	645.73	870	PALL protein - yeast (Saccha
pir2:C84527	7.00	92.10	654.89	744	probable receptor-like protein	pir2:D96557	7.00	90.99	645.67	871	probable protein kinase [imp
pir2:E84698	7.00	92.08	654.65	747	hypothetical protein At2g29620	pir1:QRRBVD	7.00	90.97	645.53	873	VLDL receptor precursor - ra
pir1:TVVPTH	7.00	92.04	654.33	751	large T antigen - hamster polyo	pir1:IA8952	7.00	90.97	645.53	873	VLDL receptor precursor, lon
pir2:H90124	7.00	92.02	654.18	753	hypothetical protein orf753 [im	pir2:D88482	7.00	90.97	645.53	873	protein C05D11.8 [imported]
pir2:G83308	7.00	91.99	653.94	756	Similar to disease resistance p	pir2:B75514	7.00	90.97	645.53	873	penicillin-binding protein 1
pir2:D65622	7.00	91.98	653.87	757	hypothetical protein - Deinococ	pir2:S25530	7.00	90.96	645.47	874	glycoprotein B - bovine herp
pir2:T51428	7.00	91.98	653.87	757	hypothetical protein T9L3_130 -	pir2:H96557	7.00	90.95	645.40	875	probable protein kinase [imp
pir2:T38124	7.00	91.97	653.87	758	hypothetical protein SPAC2068.0	pir2:S60137	7.00	90.91	645.07	880	beta-N-acetylhexosaminidase
pir2:T15577	7.00	91.97	653.79	758	hypothetical protein C23G10.8 -	pir2:E96557	7.00	90.89	644.87	883	probable protein kinase [imp
pir2:T16800	7.00	91.97	653.79	758	hypothetical protein T05A7.6 -	pir2:T02731	7.00	90.88	644.81	884	serine/threonine-specific pr
pir2:I38593	7.00	91.96	653.71	759	fibroblast activation protein-a	pir2:D96730	7.00	90.88	644.81	884	unknown protein F5A18.20 [im
pir2:S00703	7.00	91.94	653.56	761	photosystem I protein Ai - gard	pir2:B89783	7.00	90.87	644.74	885	transporter homolog ydgh - B
pir2:T42430	7.00	91.94	653.56	761	hypothetical protein R166.5 - C	pir2:T39081	7.00	90.86	644.68	886	hypothetical protein SPAC704
pir1:H70414	7.00	91.92	653.32	764	conserved hypothetical protein	pir2:T20941	7.00	90.86	644.61	887	hypothetical protein F15A2.6
pir2:T39878	7.00	91.92	653.32	764	hypothetical protein SPAC20F10.	pir2:T00800	7.00	90.83	644.42	890	disease resistance protein h
pir2:B86597	7.00	91.90	653.17	766	CT111 hypothetical protein [imp	pir2:T40137	7.00	90.82	644.35	891	hypothetical serine rich pro
pir2:T01817	7.00	91.90	653.17	766	hypothetical protein T27D20.9 -	pir2:F97517	7.00	90.82	644.29	892	DNA topoisomerase I (omega-p
pir2:T72027	7.00	91.89	653.17	766	conserved hypothetical protein	pir2:AN2736	7.00	90.82	644.29	892	DNA topoisomerase I topA [R
pir2:I53821	7.00	91.88	653.02	768	P-selectin - rat	pir2:T01899	7.00	90.82	644.29	892	disease resistance protein R
pir2:S52684	7.00	91.88	653.02	768	probable membrane protein YDR11	pir2:H96651	7.00	90.81	644.22	893	protein T3P18.19 [imported]
pir2:A35848	7.00	91.87	652.94	769	competence regulatory protein c	pir2:T17276	7.00	90.81	644.22	893	hypothetical protein DKF2p43
pir2:B69604	7.00	91.87	652.94	769	two-component sensor histidine	pir2:T11979	7.00	90.79	644.09	895	preprotein translocase subun
pir2:T00204	7.00	91.86	652.87	770	LDL receptor related protein 1c	pir2:S43074	7.00	90.72	643.51	896	epidermal growth factor rece
pir2:T00203	7.00	91.85	652.87	770	LDL receptor-related protein 1c	pir2:T46170	7.00	90.72	643.51	904	disease resistance-like prot
pir2:T25815	7.00	91.85	652.79	771	disease resistance protein homc	pir2:T00475	7.00	90.71	643.45	905	probable disease resistance
pir2:S25814	7.00	91.85	652.79	771	hypothetical protein YKR010c -	pir2:T41438	7.00	90.71	643.38	906	probable resistance protein



pir2:A86460 + 7.00 90.70 643.32 907 99.9k hypothetical protein T1E4  
pir2:S74860 - 7.00 90.66 643.00 912 DNA mismatch repair protein - s  
pir2:B96592 + 7.00 90.64 642.88 914 hypothetical protein T7N22.1 [i  
pir2:JC4361 - 7.00 90.61 642.62 918 scavenger receptor Cys-rich epi  
pir2:A33718 - 7.00 90.59 642.43 921 retinoblastoma protein - mouse  
pir2:IF84593 - 7.00 90.59 642.43 921 hypothetical protein At2g20810  
pir2:T03854 + 7.00 90.58 642.43 921 alkaline phosphatase synthesis  
pir2:T03854 + 7.00 90.58 642.37 922 hypothetical protein C02A12.2 -  
pir1:RBHU - 7.00 90.54 642.00 928 retinoblastoma-associated prote  
pir2:G91198 - 7.00 90.49 641.63 934 Gamma Intimin [imported] - Esch  
pir2:C86045 - 7.00 90.49 641.63 934 Intimin adherence protein [impc  
pir1:IA1193 - 7.00 90.48 641.56 935 outer membrane protein eae - Es  
pir2:TI3014 + 7.00 90.43 641.13 942 cytochrome b245 beta chain homc

seq\_name: pir2:T05322

seq\_documentation\_block:  
hypothetical protein F18F4.240 - Arabidopsis thaliana  
N:Alternate names: hypothetical protein F1C12.60  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 31-Jan-2000  
C:Accession: T05322; T04898  
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: T15408  
A:Accession: T05322  
A:Molecule type: DNA  
A:Residues: 1-1232 <BEV>  
A:Cross-references: EMBL:AL022224  
A:Experimental source: cultivar Columbia; BAC clone F1C12  
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, February 1998  
A:Reference number: T15398  
A:Accession: T04898  
A:Molecule type: DNA  
A:Residues: 1-305 <BEW>  
A:Cross-references: EMBL:AL021637  
A:Experimental source: cultivar Columbia; BAC clone F18F4  
C:Genetics:  
A:Map position: 4  
A:Introns: 863/1; 1116/1  
A:Note: F1C12.60; F18F4.240  
C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; P

alignment\_scores:  
Quality: 12.00 Length: 12  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x T05322 ..

Align seg 1/1 to: T05322 from: 1 to: 1232

2185 CTGGAACCTTGACCTCAGCCACACCAACTGACC 2220

794 LeuGluThrLeuAspLeuSerHisnGlnLeuThr 805

seq\_name: pir2:AC1328

seq\_documentation\_block:  
internalin proteins homolog lmo2027 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
J.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AC1328  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-367 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAD00105.1; PID:g16411497; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo2027

alignment\_scores:

Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AC1328 ..

Align seg 1/1 to: AC1328 from: 1 to: 367

2191 ACTTTGGACCTCAGCCACACCAACTGACC 2220

201 ThrLeuAspLeuSerHisnGlnLeuThr 210

seq\_name: pir2:S67265

seq\_documentation\_block:

hypothetical protein YOR353c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O6612  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: S67265  
R:Delius, H.; Hebling, U.; Hofmann, B.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67261  
A:Accession: S67265  
A:Molecule type: DNA  
A:Residues: 1-791 <DEL>  
A:Cross-references: EMBL:275261; NID:g1420767; PID:e252177; PFD:g1420768; GSPDB:GN000  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YOR353c  
A:Map position: 15R

alignment\_scores:

Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S67265 ..

Align seg 1/1 to: S67265 from: 1 to: 791

706 GTGCTCTCCTGAAAGATAACAATGTGACA 735

141 ValLeuSerLeuLysAspAsnValThr 150

seq\_name: pir2:C96673

seq\_documentation\_block:

gamma-tocopherol methyltransferase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96673  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar,  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: C96673  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-348 <STO>  
A:Cross-references: GB:AE005173; NID:g8099780; PIDN:AAD38271.2; GSPDB:GN00141  
C:Genetics:  
A:Gene: F13011.27  
A:Map position: 1

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x C96673 ..  
Align seg 1/1 to: C96673 from: 1 to: 348

3068 TTCAGAGTCCAAAGTTCCTCCAGCTCC 3094  
|||||  
30 PheArgSerProSerSerSerSer 38

seq\_name: p1r2:C86291

seq\_documentation\_block:  
hypothetical protein F7H2.5 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
A:Accession: C86291  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dwar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: C86291  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <STO>  
A:Cross-references: GB:AE005172; NID:g8927650; PIDN:AAP82141.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2/rev x C86291 ..  
Align seg 1/1 to: C86291 from: 1 to: 358

2317 CTCTGAGAAATACTTCGTCCAGACTCC 2291  
|||||  
268 LeuValGluAsnThrSerSerSer 276

seq\_name: p1r1:NBHUC8

seq\_documentation\_block:  
decorin precursor - human  
N:Alternate names: cartilage proteoglycan protein II; DS-PG II; PG40 core protein; pr  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 21-Jan-2000  
A:Accession: A45016; A45015; A26476; S05640  
R:Vetter, U.; Vogel, W.; Just, W.; Young, M.F.; Fisher, L.W.  
Genomics 15, 161-168, 1993  
A:Title: Human decorin gene: intron-exon junctions and chromosomal localization.  
A:Reference number: A45016; MUID:93162643  
A:Accession: A45016  
A:Molecule type: DNA  
A:Residues: 1-359 <VET>  
A:Cross-references: GB:L01125; GB:L01126; GB:L01127; GB:L01128; GB:L01129; GB:L01130;  
A:Note: sequence extracted from NCBI backbone (NCBIP:125061)  
R:Danielson, K.G.; Fazzio, A.; Cohen, I.; Cannizzaro, L.A.; Eichstetter, I.; Iozzo, R  
Genomics 15, 146-160, 1993  
A:Title: The human decorin gene: intron-exon organization, discovery of two alternati  
A:Reference number: A45015; MUID:93162642  
A:Accession: A45015  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 28-70 <DA2>  
A:Cross-references: GB:M98262  
A:Note: sequence extracted from NCBI backbone (NCBIP:125013)  
A:Accession: B45015  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 296-359 <DAN>  
A:Note: sequence extracted from NCBI backbone (NCBIP:125017)  
R:Krusius, T.; Ruoslahti, E.  
Proc. Natl. Acad. Sci. U.S.A. 83, 7683-7687, 1986  
A:Title: Primary structure of an extracellular matrix proteoglycan core protein deduc  
A:Reference number: A26476; MUID:87017013  
A:Accession: A26476  
A:Molecule type: mRNA  
A:Residues: 1-359 <KRU>  
A:Cross-references: GB:M14219; NID:g181169; PIDN:AAB00774.1; PID:g181170  
R:Roughley, P.J.; White, R.J.  
Biochem. J. 262, 823-827, 1989  
A:Title: Dermatan sulphate proteoglycans of human articular cartilage. The properties  
A:Reference number: S05639; MUID:90073579  
A:Accession: S05640  
A:Molecule type: protein  
A:Residues: 31-33,'X',35-50 <ROU>  
C:Comment: This protein binds type I collagen.  
C:Genetics:  
A:Gene: GDB:DCN  
A:Cross-references: GDB:119839; OMIM:125255  
A:Map position: 12q21.3-12q23  
A:Introns: 71/1; 108/3; 180/1; 218/1; 249/2; 295/3  
A:Note: the first two introns occur before the initiator codon  
C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
C:Keywords: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; dup  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-30/Domain: propeptide #status predicted <PRO>  
F:31-359/Product: decorin #status predicted <MPT>  
F:48-72/Domain: proteoglycan amino-terminal homology <PAH>  
F:82-105/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:106-129/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:130-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:151-174/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:175-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:201-221/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:222-245/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:246-269/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:270-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:308-359/Domain: proteoglycan carboxyl-terminal homology <PRC>  
F:34/Binding site: dermatan sulfate (Ser) (covalent) #status experimental  
F:189,325/Binding site: dermatan sulfate (Ser) (covalent) #status predicted  
F:211,262,303/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-202-054-2 x NBHUC8 ..
  Align seg 1/1 to: NBHUC8 from: 1 to: 359
    508 ATACCGCAGGCGCTCCCGCTAGCTTA 534
    |||||
    215 IleProGlnGlyLeuProSerLeu 223

seq_name: pir2:S06280

seq_documentation_block:
  decorin precursor - bovine
  N:Alternate names: dermatan sulfate proteoglycan II; proteoglycan core protein II
  C:Species: Bos primigenius taurus (cattle)
  C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Aug-1999
  C:Accession: S06280; B31430; A26545; A20935
  R:Day, A.A.; McQuillan, C.I.; Termino, J.D.; Young, M.R.
  Biochem. J. 248, 801-805, 1987
  A:Title: Molecular cloning and sequence analysis of the cDNA for small proteoglycan II
  A:Reference number: S06280; MUID:88133946
  A:Accession: S06280
  A:Molecule type: mRNA
  A:Residues: 1-360 <DAY>
  A:Cross-references: EMBL:Y00712; NID:g618; PIDN:CAA68702.1; PID:g619
  A:Experimental source: bone
  R:Choi, H.U.; Johnson, T.L.; Pal, S.; Tang, L.H.; Rosenberg, L.; Neame, P.J.
  J. Biol. Chem. 264, 2876-2884, 1989
  A:Title: Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII, from
  A:Reference number: A31430; MUID:89123388
  A:Accession: B31430
  A:Molecule type: protein
  A:Residues: 31-33, 'X', 35-54 <CHO>
  A:Experimental source: cartilage; fetal skin
  R:Cosser, L.; Rosenberg, L.C.; van der Rest, M.; Poole, A.R.
  J. Biol. Chem. 262, 3809-3812, 1987
  A:Title: The dermatan sulfate proteoglycans of bovine sclera and their relationship to
  A:Reference number: A26545; MUID:87137687
  A:Accession: A26545
  A:Molecule type: protein
  A:Residues: 31-50 <COS>
  A:Experimental source: sclera
  R:Pearson, C.H.; Winterbottom, N.; Fackre, D.S.; Scott, P.G.; Carpenter, M.R.
  J. Biol. Chem. 258, 15101-15104, 1983
  A:Reference number: A20935; MUID:84087911
  A:Accession: A20935
  A:Molecule type: protein
  A:Residues: 31-54 <PEA>
  A:Experimental source: skin
  R:Chopra, R.K.; Pearson, C.H.; Pringle, G.A.; Fackre, D.S.; Scott, P.G.
  Biochem. J. 232, 277-279, 1985
  A:Title: Dermatan sulphate is located on serine-4 of bovine skin proteodermatan sulphate
  nes around glycosylation sites in different proteoglycans.
  A:Reference number: A44700; MUID:86103195
  A:Contents: annotation; glycosylation
  C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology <PAH>
  C:Keywords: chondroitin sulfate proteoglycan; collagen binding; dermatan sulfate; extrac
  F:1-15/Domain: signal sequence #status predicted <SIG>
  F:16-30/Domain: propeptide #status predicted <PRO>
  F:31-360/Product: decorin #status predicted <NAR>
  F:83-106/Domain: proteoglycan amino-terminal homology <PAH>
  F:107-130/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
  F:131-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
  F:152-175/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
  F:176-199/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
  F:202-222/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
  F:223-246/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
  F:247-270/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
  F:271-293/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
  F:294-308/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
  F:309-360/Domain: proteoglycan carboxyl-terminal homology <PCH>

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
  US-09-202-054-2 x I47020 ..
  Align seg 1/1 to: I47020 from: 1 to: 360
    508 ATACCGCAGGCGCTCCCGCTAGCTTA 534
    |||||
    216 IleProGlnGlyLeuProSerLeu 224

seq_name: pir1:NBHUIA

seq_documentation_block:
  platelet glycoprotein Ib alpha chain precursor - human
```



N:Alternate names: membrane glycoprotein Ib alpha chain  
N:Contains: glycosialicin  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 22-Jun-1999  
C:Accession: A941174; A60435; A941173; S16945; I55355; A27075; A27102  
R:Lopez, J.A.; Chung, D.W.; Fujikawa, K.; Hegen, F.S.; Papayannopoulou, T.; Roth, G.J.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5615-5619, 1987  
A:Title: Cloning of the alpha-chain of human platelet glycoprotein Ib: a transmembrane protein  
A:Reference number: A941174; MUID:87289655  
A:Accession: A941174  
A:Molecule type: mRNA  
A:Residues: 1-626 <LQP>  
A:Cross-references: GB:J02940; NID:G183499; PIDN:AAA52595.1; PID:G306793  
R:Wicki, A.N.; Walz, A.; Gerber-Huber, S.N.; Wenger, R.H.; Vornhagen, R.; Clemetson, K.J.  
Thromb. Haemost. 61, 448-453, 1989  
A:Title: Isolation and characterization of human blood platelet mRNA and construction of a cDNA library  
A:Reference number: A60435; MUID:90020160  
A:Accession: A60435  
A:Molecule type: mRNA  
R:Titanai, K.; Takio, K.; Handa, M.; Ruggeri, Z.M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5610-5614, 1987  
A:Title: Amino acid sequence of the von Willebrand factor-binding domain of platelet membrane glycoprotein Ib  
A:Reference number: A941173; MUID:87289654  
A:Accession: A941173  
A:Molecule type: protein  
A:Residues: 17-315 <NTT>  
R:Hess, D.; Schaller, J.; Rickli, E.E.; Clemetson, K.J.  
Eur. J. Biochem. 199, 389-393, 1991  
A:Title: Identification of the disulphide bonds in human platelet glycosialicin.  
A:Reference number: S16945; MUID:91301149  
A:Accession: S16945  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 224-227;262-270;277-282 <HES>  
R:Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.  
J. Biol. Chem. 267, 10055-10061, 1992  
A:Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of tandem repeats  
A:Reference number: I55355; MUID:92250564  
A:Accession: I55355  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 412-427 <RES>  
A:Cross-references: GB:S34436; NID:G249176; PIDN:AAB22152.1; PID:G249177  
A>Note: variant D  
C:Comment: Glycoprotein Ib (GPb), a surface membrane protein of platelets, participates in platelet activation apparently involves disruption of the macromolecular complex of binding sites for von Willebrand factor and thrombin (the latter site with von Willebrand factor). Glycosialicin, which is approximately coextensive with the extracellular part of the protein, is a major component of the protein.  
C:Genetics:  
A:Gene: GDB:GP1BA; GP1B  
A:Cross-references: GDB:118806; OMIM:231200  
A:Map position: 1p11p12  
C:Complex: heterodimer with platelet glycoprotein Ib beta chain (NBHUIB)  
C:Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein repeat  
C:Keywords: blood coagulation; duplication; glycoprotein; platelet membrane; tandem repeat  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-626/Product: platelet glycoprotein Ib alpha chain #status predicted <MPT>  
F:48-71/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:72-93/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:94-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:117-140/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:141-164/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:165-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:189-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:379-430/Region: proline/threonine-rich 9-residue repeats  
F:502-540/Domain: transmembrane #status predicted <TRM>  
F:541-626/Domain: intracellular #status predicted <INT>  
F:37,175/Binding site: carboxydrate (Asn) (covalent) #status experimental  
F:308/Binding site: carbohydrate (Thr) (covalent) #status experimental

alignment\_scores: Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x NBHUIA ..  
Align seg 1/1 to: NBHUIA from: 1 to: 626  
2191 ACTTTGGACCTCAGCCACCAACCACTG 2217  
|||||  
97 ThrLeuAspLeuSerHisAsnGlnLeu 105  
seq\_name: pirl:A29944  
seq\_documentation\_block:  
chaoptin precursor - fruit fly (Drosophila melanogaster)  
N:Alternate names: photoreceptor cell-specific membrane protein  
C:Species: Drosophila melanogaster  
C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 22-Jun-1999  
C:Accession: A29944; A21123  
R:Reinke, R.; Krantz, D.E.; Yen, D.; Zipursky, S.L.  
Cell 52, 291-301, 1988  
A:Title: Chaoptin, a cell surface glycoprotein required for Drosophila photoreceptor development  
A:Reference number: A29944; MUID:88135762  
A:Accession: A29944  
A:Molecule type: DNA  
A:Residues: 1-1134 <REI>  
A:Cross-references: GB:M19008; GB:M19009; GB:M19010; GB:M19011; GB:M19012; GB:M19013;  
R:Zipursky, S.L.; Venkatesh, T.R.; Teplov, D.B.; Benzer, S.  
Cell 36, 15-26, 1984  
A:Title: Neuronal development in the Drosophila retina: monoclonal antibodies as mole  
A:Reference number: A21123; MUID:84106810  
A:Accession: A21123  
A:Molecule type: protein  
A:Residues: 31-43,'HX','46-49,'H' <ZIP>  
C:Genetics:  
A:Gene: FlyBase:cbp  
A:Cross-references: FlyBase:FBgn0000313  
A:Introns: 1/3 80/3; 318/3; 377/2 422/2; 702/1 745/3; 831/2; 998/2  
C:Superfamily: chaoptin; leucine-rich alpha-2-glycoprotein repeat homology  
C:Keywords: cell adhesion; glycoprotein; membrane protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-1134/Product: chaoptin #status predicted <MAT>  
F:80-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <  
F:103-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:128-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:152-175/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:177-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:201-224/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:226-249/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:250-273/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>  
F:279-302/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
F:303-325/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>  
F:326-349/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
F:351-374/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
F:375-399/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>  
F:401-424/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
F:428-451/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>  
F:453-476/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>  
F:477-500/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>  
F:502-526/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>  
F:527-550/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>  
F:551-574/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR20>  
F:577-600/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR21>  
F:601-624/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR22>  
F:625-648/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR23>  
F:649-672/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR24>  
F:673-696/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR25>  
F:708-731/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR26>  
F:733-756/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR27>

F:757-780/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR28>  
F:781-804/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR29>  
F:805-827/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR30>  
F:828-851/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR31>  
F:854-877/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR32>  
F:879-902/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR33>  
F:903-926/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR34>  
F:928-948/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR35>  
F:949-972/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR36>  
F:973-995/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR37>  
F:996-1019/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR38>  
F:1021-1044/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR39>  
F:1056-1080/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR40>

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x A29944 ..

Align seg 1/1 to: A29944 from: 1 to: 1134

835 TTCAAAATCTTGACCTAACTGGAAT 861

||||| 835 LeuGlnLeuAspLeuSerGlyAsn 360

seq\_name: pir2:A49674

seq\_documentation\_block:

flightless-I homolog - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000

C:Accession: A49674

R:Campbell, H.D.; Schimanski, T.; Claudianos, C.; Ozsarac, N.; Kasprzak, A.B.; Cotseil,

Proc. Natl. Acad. Sci. U.S.A. 90, 11386-11390, 1993

A:Title: The Drosophila melanogaster flightless-I gene involved in gastrulation and musc

s and humans.

A:Reference number: A49674; MUID:94068608

A:Accession: A49674

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1268 <RES>

A:Cross-references: EMBL:U01184; NID:9440176; PIDN:AAC03568.1; PID:9440177

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; gelsolin repeat homolo

F:498-838/Domain: gelsolin repeat homology <GEL1>

F:904-1261/Domain: gelsolin repeat homology <GEL2>

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x A49674 ..

Align seg 1/1 to: A49674 from: 1 to: 1268

2194 TTGGACCTCAGCCACCACTGACC 2220

||||| 108 LeuAspLeuSerHisGlnLeuThr 116

seq\_name: pir2:S56221

seq\_documentation\_block:

hypothetical protein YFL033c - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 02-Sep-1995 #sequence\_revision 19-Oct-1995 #text\_change 29-Oct-1999

C:Accession: S56221

R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasano

submitted to the EMBL Data Library, May 1995  
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces  
A:Reference number: S56186  
A:Accession: S56221  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1770 <MUR>  
A:Cross-references: EMBL:D50617; NID:9836685; PIDN:BAA09206.1; PID:d1009846; PID:9836  
C:Genetics:  
A:Gene: SGD:RIM15  
A:Cross-references: SGD:S0001861; MIPS:YFL033c  
A:Map position: 6L

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S56221 ..

Align seg 1/1 to: S56221 from: 1 to: 1770

2107 CCAATCTAAAGAACTCTCTTTGGGC 2133

||||| 1397 ProAsnLeuLysAsnLeuSerLeuAla 1405

seq\_name: pir2:C96615

seq\_documentation\_block:

hypothetical protein T18124.10 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: C96615

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: C96615

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1784 <STO>

A:Cross-references: GB:AE005173; NID:g11038494; PIDN:ANG27771.1; GSPDB:GN00141

C:Genetics:

A:Gene: T18124.10

A:Map position: 1

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x C96615 ..

Align seg 1/1 to: C96615 from: 1 to: 1784

1030 CTCAGCACTGGATCTGCCAAAC 1056

||||| 226 LeuGlnLeuAspLeuSerGlnAsn 234

seq\_name: pir2:T21312

seq\_documentation\_block:

hypothetical protein F23D12.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T21312; T24907  
R:Barlow, K.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19404  
A:Accession: T21312  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2287 <W12>  
A:Cross-references: EMBL:Z71186; PIDN:CAA94917.1; GSPDB:GN00028; CESP:F23D12.2  
A:Experimental source: clone F23D12  
R:Barlow, K.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19952  
A:Accession: T24907  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2287 <W12>  
A:Cross-references: EMBL:Z70687; NID:g1256502; PIDN:CAA94618.1; GSPDB:GN00028; CESP:F23D12.2  
A:Experimental source: clone T14C1  
C:Genetics:  
A:Gene: CESP:F23D12.2  
A:Map position: X  
A:Introns: 75/3; 141/3; 165/3; 409/3; 497/3; 572/1; 583/3; 686/3; 773/3; 916/3; 945/3; 1  
  
alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2/rev x T21312 ..  
  
Align seg 1/1 to: T21312 from: 1 to: 2287  
  
1636 AAGAAAGATGCTGAAATCAGAGGACT 1610  
|||||  
205 LysLysAspAlaGluasnGlnargThr 213  
  
seq\_name: plr2:S68976  
  
seq\_documentation\_block:  
tetrahydromethanopterin S-methyltransferase (EC 2.1.1.86) chain G mtrG [validated] - Met  
N:Alternate names: N5-methyltetrahydromethanopterin--coenzyme M methyltransferase mtrG 1  
C:Species: Methanobacterium thermoautotrophicum  
A:Variety: strain Marburg, DSM 2133  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 02-Sep-2000  
C:Accession: S68976  
R:Harms, U.; Weiss, D.S.; Gaertner, P.; Linder, D.; Thauer, R.K.  
Eur. J. Biochem. 228, 640-648, 1995  
A:Title: The energy conserving N(5)-methyltetrahydromethanopterin:coenzyme M methyltrans  
A:Reference number: S68974; MUID:95255265  
A:Accession: S68976  
A:Molecule type: DNA  
A:Residues: 1-86 <HAR>  
A:Cross-references: EMBL:X84219; NID:g668976; PIDN:CAA59002.1; PID:g668978  
A:Experimental source: strain Marburg, DSM 2133  
C:Genetics:  
A:Gene: mtrG  
C:Complex: membrane-associated complex; heterooctamer of chains A (see PIR:S38369), B (s  
IR:S68977)  
C:Function:  
A:Description: catalyzes the reaction of N(5)-methyltetrahydromethanopterin and coenzyme  
A:Pathway: methanogenesis  
C:Superfamily: Methanobacterium tetrahydromethanopterin S-methyltransferase chain G  
C:Keywords: membrane-associated complex; methanogenesis; methyltransferase  
  
alignment\_scores:  
Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2/rev x S68976 ..  
  
Align seg 1/1 to: S68976 from: 1 to: 86  
  
3078 GGACTTCTGAAAGGGCTTCTCAAG 3055  
|||||  
74 GlyLeuLeuLysGlyLeuLeuLys 81  
  
seq\_name: pir2:B69021  
  
seq\_documentation\_block:  
tetrahydromethanopterin S-methyltransferase (EC 2.1.1.86) chain G MTH1157 [similarity  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 02-Sep-2000  
C:Accession: B69021  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan,  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
A:Reference number: A69000; MUID:98037514  
A:Accession: B69021  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-86 <MTH>  
A:Cross-references: GB:AE000885; GB:AE000666; NID:g2622256; PIDN:AAB85646.1; PID:g262  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1157  
C:Superfamily: Methanobacterium tetrahydromethanopterin S-methyltransferase chain G  
C:Keywords: methyltransferase  
  
alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2/rev x B69021 ..  
  
Align seg 1/1 to: B69021 from: 1 to: 86  
  
3078 GGACTTCTGAAAGGGCTTCTCAAG 3055  
|||||  
74 GlyLeuLeuLysGlyLeuLeuLys 81  
  
seq\_name: pir2:C70010  
  
seq\_documentation\_block:  
Na+/H+ antiporter homolog yufv - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: C70010  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; MUID:98044033

A:Accession: C70010

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-113 <KUN>

A:Cross-references: GB:Z99120; GB:AL009126; NID:G2635613; PIDN:CABI5151.1; PID:G2635658

A:Experimental source: strain 168

C:Genetics:

A:Gene: yufv

C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1887

alignment\_scores:

Quality:	8.00	Length:	8
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-202-054-2 x C70010 ..

Align seg 1/1 to: C70010 from: 1 to: 113

630 CTGTTATTATCGAATCCTGTGTA 653

|||||

19 LeuLeuSerLySerLeuLeu 26

seq\_name: pir2:A95285

seq\_documentation\_block:

hypothetical protein SMA0343 [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C:Accession: A95285

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Faim, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95282; MUID:21396509; PMID:11481432

A:Accession: A95285

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <KUR>

A:Cross-references: GB:AB006469; PIDN:AAK64843.1; PID:G14523256; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 688-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMA0343

A:Genome: plasmid

alignment\_scores:

Quality:	8.00	Length:	8
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-202-054-2/rev x A95285 ..

Align seg 1/1 to: A95285 from: 1 to: 144

517 CTTGCGGTATCTAGTACTGCTGT 494

|||||

61 ProAlaValSerLeuValAlaGly 68

seq\_name: pir2:C97180

seq\_documentation\_block:

uncharacterized protein Yih2 family [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001

C:Accession: C97180

R:Nolling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C97180

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-149 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK80230.1; PID:G15025277; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC2273

C:Superfamily: conserved hypothetical protein HI0670

alignment\_scores:

Quality:	8.00	Length:	8
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-202-054-2/rev x C97180 ..

Align seg 1/1 to: C97180 from: 1 to: 149

773 AGTCTGTTAAAGTAGTCGCAA 750

|||||

11 SerSerVallySvalaspGlyLys 18

seq\_name: pir2:A03864

seq\_documentation\_block:

hypothetical 17.7K protein - human adenovirus 2

C:Species: Mastadenovirus h2 (human adenovirus 2)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Sep-1999

C:Accession: F92351; F92352; A03864

R:Gingeras, T.R.; Sclaky, D.; Gellinas,

J. Biol. Chem. 257, 13475-13491, 1982

A:Title: Nucleotide sequences from the adenovirus-2 genome.

A:Reference number: A92351; MUID:83056843

A:Accession: F92351

A:Molecule type: DNA

A:Residues: 1-168 <GIN>

R:Alestrom, P.; Akusjarvi, G.; Pettersson, M.; Pettersson, U.

J. Biol. Chem. 257, 13492-13498, 1982

A:Title: DNA sequence analysis of the region encoding the terminal protein and the

A:Reference number: A92352; MUID:83056844

A:Accession: F92352

A:Molecule type: DNA

A:Residues: 1-168 <ALE>

C:Superfamily: human adenovirus 2 hypothetical 17.7K protein

alignment\_scores:

Quality:	8.00	Length:	8
Ratio:	1.000	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment\_block:

US-09-202-054-2 x A03864 ..

Align seg 1/1 to: A03864 from: 1 to: 168

3074 ACTCCAAGTCTCTCCAGCTCCGGA 3097

|||||

29 SerProSerSerSerSerGly 36

seq\_name: plr2:D83323

seq\_documentation\_block:  
probable acetyltransferase PA2578 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83323  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:2043737  
A:Accession: D83323  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-186 <STO>  
A:Cross-references: GB:AE004686; GB:AE004091; NID:g9948636; PIDN:AAG05966.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2578

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x D83323 ..  
Align seg 1/1 to: D83323 from: 1 to: 186

1029 ACTCCAGGAACTGGATCTGTCCCA 1052  
|||||  
3 ThrProGlyThrGlySerValPro 10

seq\_name: plr2:T02792

seq\_documentation\_block:  
hypothetical protein L549.4 [imported] - Leishmania major (strain Friedlin)  
C:Species: Leishmania major  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 19-May-2000  
C:Accession: D81455; T02792  
R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;  
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999  
A>Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c  
A:Reference number: A81455; MUID:99178987  
A:Accession: D81455  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-189 <PYL>  
A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24616.1; PID:g2978453; GSPDB:GN00  
A:Experimental source: strain MHOM/IL/81/Friedlin  
C:Genetics:  
A:Gene: L549.4  
A:Map position: 1  
A:Note: L549.4  
C:Superfamily: Leishmania major hypothetical protein L549.4

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x T02792 ..  
Align seg 1/1 to: T02792 from: 1 to: 189

3071 AGAAGTCCAAGTTCTCCAGCTCC 3094

|||||  
31 ArgSerProSerSerSerSerSer 38

seq\_name: plr2:A13239

seq\_documentation\_block:  
conserved hypothetical protein Atu6099 [imported] - Agrobacterium tumefaciens (strain  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: A13239  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCCI  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: A13239  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-200 <KUR>  
A:Cross-references: GB:AE008690; PIDN:AAL46335.1; PID:g17744123; GSPDB:GN00189  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu6099  
A:Genome: plasmid

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x A13239 ..  
Align seg 1/1 to: A13239 from: 1 to: 200

1167 TCTATCACAGCATTTTCTTCACT 1190  
|||||  
79 SerIleThrSerIlePhePheThr 86

seq\_name: plr1:S73914

seq\_documentation\_block:  
MGI05 homolog K04\_orf202 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: S73914  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkle, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia  
A:Reference number: S73327; MUID:97105885  
A:Accession: S73914  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-202 <HIM>  
A:Cross-references: EMBL:AE000057; GB:U00089; NID:g1674279; PIDN:AAB96236.1; PID:g167  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Genetic code: SGC3  
C:Superfamily: hypothetical protein ybbp

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S73914 ..

Align seg 1/1 to: S73914 from: 1 to: 202

1286 TTCTTGATCTTGGCACTAACTTTA 1309  
|||||  
16 PheUuIfuLeuAlaLeuThrLeu 23

seq\_name: pir2:T01040

seq\_documentation\_block:  
hypothetical protein YUP8H12R.23 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 24-Nov-1999  
C:Accession: T01040  
R:Theologis, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwack, P.; Davis, R.W.  
submitted to the EMBL Data Library, May 1998  
A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.  
A:Reference number: Z14227  
A:Accession: T01040  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-204 <TH>  
A:Cross-references: EMBL:AC002986; NID:g2494106; PID:g3152565; GSPDB:GN00059; ATSP:YUP8H12R.23  
C:Genetics:  
A:Gene: ATSP:YUP8H12R.23  
A:Map position: 1  
A:Introns: 71/3  
C:Superfamily: Arabidopsis thaliana hypothetical protein YUP8H12R.23

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x T01040 ..

Align seg 1/1 to: T01040 from: 1 to: 204

690 CTTGACAAAGTTAAAGTGCTCTC 713  
|||||  
118 LeuAspLysValLysSerAlaLeu 125

seq\_name: pir2:F70469

seq\_documentation\_block:  
enolase-phosphatase E-1 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 18-Aug-2000  
C:Accession: F70469  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oviatt, D.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: F70469  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-223 <AQF>  
A:Cross-references: GB:AE000766; NID:g2984216; PID:AA07754.1; PID:g2984225; GB:AE000655  
C:Experimental source: strain VF5  
C:Genetics:  
A:Gene: masA  
C:Superfamily: Klebsiella oxytoca methionine salvage pathway enzyme E-1

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x F70469 ..

Align seg 1/1 to: F70469 from: 1 to: 223

2078 AAGAACTTAGGAATTTTAGAG 2055  
|||||  
29 LysLysLeuArgGluPheLeuGlu 36

seq\_name: pir2:T37467

seq\_documentation\_block:  
ribosomal protein s4e - Thermoplasma acidophilum  
C:Species: Thermoplasma acidophilum  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T37467  
R:Thomas, N.A.; Jarrell, K.F.  
submitted to the EMBL Data Library, May 1996  
A:Description: Nucleotide sequence of the ribosomal protein genes s4e and L5 in the a  
A:Reference number: Z21705  
A:Accession: T37467  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-237 <THO>  
A:Cross-references: EMBL:U57643; PIDN:AA02244.1  
C:Genetics:  
A:Gene: rps4e  
C:Superfamily: rat ribosomal protein S4

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2/rev x T37467 ..

Align seg 1/1 to: T37467 from: 1 to: 237

767 GTTAAAGTAGTGCAAAACAGTA 744  
|||||  
68 ValLysValAspGlyLysThrVal 75

seq\_name: pir2:S05572

seq\_documentation\_block:  
hypothetical protein 238 - Streptomyces griseus  
C:Species: Streptomyces griseus  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 22-Oct-1999  
C:Accession: S05572  
R:Kobayashi, T.; Takao, M.; Oikawa, A.; Yasui, A.  
Nucleic Acids Res. 17, 4731-4744, 1989  
A:Title: Molecular characterization of a gene encoding a photolase from Streptomyces  
A:Reference number: S05572; MUID:89315214  
A:Accession: S05572  
A:Molecule type: DNA  
A:Residues: 1-238 <KOB>  
A:Cross-references: EMBL:X15060; NID:g47081; PIDN:CAA33160.1; PID:g47082

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x S05572 ..

Align seg 1/1 to: S05572 from: 1 to: 238

1932 CATGGTCCTGCTGGTGAGAGA 1909  
|||||

72 HisGlyProAlaGlyGlyArg 79

seq\_name: pir2:S10005

seq\_documentation\_block:  
hypothetical protein 5 - fowl adenovirus 1  
C:Species: Avidadenovirus galli (fowl adenovirus 1, CELO)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 20-Apr-2000  
C:Accession: S10005  
R:Akopian, T.A.; Kruglyak, V.A.; Rivkina, M.B.; Naroditsky, B.S.; Tikhonenko, T.I.  
Nucleic Acids Res. 18, 2825, 1990  
A:Title: Sequence of an avian adenovirus (CELO) DNA fragment (0-11.2%).  
A:Reference number: S10004; MUID:90251474  
A:Accession: S10005  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-283 <AKO>  
A:Cross-references: EMBL:X17217; NID:g58537; PIDN:CAA35087.1; PID:g58539  
C:Superfamily: fowl adenovirus 1 hypothetical protein 5

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S10005 ..

Align seg 1/1 to: S10005 from: 1 to: 283

286 ACGAACCTCACCTCACCATTAAC 309  
|||||  
222 ThrAsnLeuThrLeuThrIleAsn 229

seq\_name: pir2:AI0810

seq\_documentation\_block:  
pyridoxal kinase (EC 2.7.1.35) - Salmonella enterica subsp. enterica serovar Typhi (stra  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: This species has also been called Salmonella typhi  
C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AI0810  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AI0810  
A:Molecule type: DNA  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-288 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD07667.1; PID:g16503653; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY2672  
C:Keywords: phosphotransferase

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x AI0810 ..

Align seg 1/1 to: AI0810 from: 1 to: 288

730 GTCACAGCCGTCCTACTGTTTG 753  
|||||  
51 ValThrAlaValProThrValLeu 58

seq\_name: pir2:C83972

seq\_documentation\_block:  
2-dehydropantoate 2-reductase apbA [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: C83972  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: AB3650; MUID:20512582; PMID:11058132  
A:Accession: C83972  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <STO>  
A:Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06298.1; GSPDB:  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: apbA

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x C83972 ..

Align seg 1/1 to: C83972 from: 1 to: 308

82 TTAGTCTCTTCCAAATGAATG 59  
|||||  
101 LeuValPhePheGlnAsnGlyMet 108

seq\_name: pir2:F86295

seq\_documentation\_block:  
T24D18.18 protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001  
C:Accession: F86295  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: AB6141; MUID:21016719  
A:Accession: F86295  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313 <STO>  
A:Cross-references: GB:AE005172; NID:g6587814; PIDN:AAF18505.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x F86295 ..

Align seg 1/1 to: F86295 from: 1 to: 313

82 TTACTCTTCCTCCAAATGAATG 59  
|||||  
118 LeuValPhePheGlnAsnGlyMet 125

seq\_name: pir2:T49908

seq\_documentation\_block:  
hypothetical protein T24H18.110 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49908  
R:Bevan, M.; Robben, J.; Grymonprez, B.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Rudd,  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z25024  
A:Accession: T49908  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-371 <BEV>  
A:Cross-references: EMBL:AL353013; GSPDB:GN00063; ATSP:T24H18.110  
A:Experimental source: cultivar Columbia; BAC clone T24H18  
C:Genetics:  
A:Gene: ATSP:T24H18.110  
A:Map position: 5

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x T49908 ..

Align seg 1/1 to: T49908 from: 1 to: 371

1648 CTGAATCTGCAGAAATTCATT 1671

|||||

282 LeuAsnLeuSerGlyAsnLeuIle 289

seq\_name: pir2:D71175

seq\_documentation\_block:  
hypothetical protein PH0600 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C:Accession: D71175  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yanamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137  
A:Accession: D71175  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-392 <KAW>  
A:Cross-references: GB:AP000002; NID:g3236129; PIDN:BAA29689.1; PID:g3257006  
A:Experimental source: strain O73  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0600  
A:Superfamily: CBS homology  
F:217-265/Domain: CBS homology <CBS>

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x D71175 ..

Align seg 1/1 to: D71175 from: 1 to: 392

2950 GTGATGACAGCAAGTATGCAAG 2973

|||||

9 ValMetThrAspLysTyrAlaLys 16

seq\_name: pir2:H96536

seq\_documentation\_block:  
hypothetical protein F2J10.8 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: H96536  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: H96536  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-460 <STO>  
A:Cross-references: GB:AE005173; NID:g8569096; PIDN:AAF76441.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F2J10.8  
A:Map position: 1

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x H96536 ..

Align seg 1/1 to: H96536 from: 1 to: 460

3071 AGAAGTCCAAAGTTCCTCCAGCTCC 3094

|||||

20 ArgSerProSerSerSerSer 27

seq\_name: pir2:C64119

seq\_documentation\_block:  
ApGlucose--starch(bacterial glycogen) glucosyltransferase (EC 2.4.1.21) - Haemophilu  
C:Species: Haemophilus influenzae  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
C:Accession: C64119  
R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage  
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: C64119  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <TIGR>  
A:Cross-references: GB:U32815; GB:L42023; NID:g1574818; PIDN:AAC23007.1; PID:g1574823  
C:Superfamily: starch synthase  
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0



Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2/rev x C64119 ..  
Align seg 1/1 to: C64119 from: 1 to: 476

1524 TTGTGTTTGAATCTGCAACTCCT 1501  
|||||  
421 PheValPheGluSerAlaThrPro 428

seq\_name: pir2:T51702

## seq\_documentation\_block:

amidophosphoribosyltransferase (EC 2.4.2.14) [similarity] - Lactococcus lactis  
N:Alternate names: phosphoribosylpyrophosphate amidotransferase  
C:Species: Lactococcus lactis  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 03-Aug-2001  
C:Accession: T51702  
R:Peitonen, T.; Mantasala, P.  
Mol. Gen. Genet. 261, 31-41, 1999  
A:Title: Isolation and characterization of a purC(orf)QLF operon from Lactococcus lactis  
A:Reference number: 225429; MUID:99168765  
A:Accession: T51702  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-506 <PEL>  
A:Cross-references: EMBL:U64311; PIDN:AAD12627.1  
A:Experimental source: strain MGL614  
C:Genetics:  
A:Gene: purF  
C:Superfamily: amidophosphoribosyltransferase  
C:Keywords: glycosyltransferase; pentosyltransferase; purine nucleotide biosynthesis  
F:46/Active site: Cys #status predicted

## alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x T51702 ..  
Align seg 1/1 to: T51702 from: 1 to: 506

2125 TCCTGGCCAAATAATGGCTCAAA 2148  
|||||  
499 SerLeuAlaLysAsnGlyLeuLys 506

seq\_name: pir2:B69205

## seq\_documentation\_block:

cobyrinic acid synthase - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
C:Accession: B69205  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.;  
Kl, S.; Church, G.W.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514  
A:Accession: B69205  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-514 <MTH>  
A:Cross-references: GB:AE000857; GB:AE000666; NID:g2621876; PIDN:AAB85289.1; PID:g262187  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH787  
A:Start codon: GTG  
C:Superfamily: probable cobyrinic acid synthase

## alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2/rev x B69205 ..  
Align seg 1/1 to: B69205 from: 1 to: 514

194 TCACAGGCGACAGTCTTTAGGAAC 171  
|||||  
387 SerGlnGlyArgValLeuGlyAsn 394

seq\_name: pir2:T32012

## seq\_documentation\_block:

hypothetical protein K02F6.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T32012  
R:Ledwith, J.; Wohlmann, P.; Rohlfing, T.  
submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of C. elegans cosmid K02F6.  
A:Reference number: Z21112  
A:Accession: T32012  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-528 <LED>  
A:Cross-references: EMBL:AF016670; PIDN:AAB66108.1; GSPDB:GN000020; CESP:K02F6.7  
A:Experimental source: strain Bristol N2; clone K02F6  
C:Genetics:  
A:Gene: CESP:K02F6.7  
A:Map position: 2  
A:Introns: 27/3; 73/3; 115/2; 210/1; 250/3; 298/3; 354/3; 474/2

## alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2/rev x T32012 ..  
Align seg 1/1 to: T32012 from: 1 to: 528

2346 GAGATCCAGATATCGCAACTGGAA 2323  
|||||  
504 GluIleGlnIleSerGlnLeuGlu 511

seq\_name: pir2:T00824

## seq\_documentation\_block:

probable thioredoxin reductase At2g41680 [imported] - Arabidopsis thaliana  
N:Alternate names: thioredoxin reductase homolog T32C6.20  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
C:Accession: T00824; G84844  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K  
submitted to the EMBL Data Library, November 1997  
A:Description: Arabidopsis thaliana chromosome II BAC T32G6 genomic sequence.  
A:Reference number: Z14163  
A:Accession: T00824  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-535 <ROU>  
A:Cross-references: EMBL:AC002510; NID:g2618683; PID:g2618704  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487  
A:Accession: G84844  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-535 <STO>  
A:Cross-references: GB:AE002093; NID:g2618704; PIDN:AAB84351.1; GSPDB:GN00139  
C:Genetics:

A:Gene: T32G6.20; At2g41680

A:Map position: 2

A:Introns: 80/1; 149/2; 184/3; 273/2; 366/3; 402/3; 431/3; 466/3

C:Superfamily: thioredoxin reductase homology

F:80-391/Domain: thioredoxin reductase homology <TRXB>

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x T00824 ..

Align seg 1/1 to: T00824 from: 1 to: 535

3074 AGTCCAGTTCCTCCAGCTCCGGA 3097

|||||

72 SerProSerSerSerGly 79

seq\_name: pir2:S26857

seq\_documentation\_block:

isocitrate lyase (EC 4.1.3.1) - *Emeritella nidulans*

N:Alternate names: isocitrate; isocitrate; isocitrate

C:Species: *Emeritella nidulans*, *Aspergillus nidulans*

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 23-Mar-2001

C:Accession: S26857; S22055

R:Galney, L.D.S.; Connerton, I.F.; Lewis, E.H.; Turner, G.; Ballance, D.J.

Curr. Genet. 21, 43-47, 1992

A:Title: Characterization of the glyoxysomal isocitrate lyase genes of *Aspergillus nidulans*

A:Reference number: S26857; MUID:92136435

A:Accession: S26857

A:Molecule type: DNA

A:Residues: 1-537 <GAT>

A:Cross-references: EMBL:X62696

R:Connerton, I.F.

submitted to the EMBL Data Library, October 1991

A:Description: Characterization of the glyoxysomal isocitrate lyase genes of *Aspergillus*

A:Reference number: S22055

A:Accession: S22055

A:Molecule type: DNA

A:Residues: 1-66,68-537 <CON>

A:Cross-references: EMBL:X62696; NID:g2316; PIDN:CAAA44572.1; PID:g2317

C:Genetics:

A:Gene: acud

A:Introns: 67/3; 119/3

C:Superfamily: isocitrate lyase

C:Keywords: carbon-carbon lyase; glyoxylate bypass; glyoxysome; oxo-acid-lyase

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S26857 ..

Align seg 1/1 to: S26857 from: 1 to: 537

1513 TTCAAAACAAAGAGGCTTCTTC 1536

|||||  
66 PhelysAsnLysGluAlaSerPhe 73

seq\_name: pir2:S44287

seq\_documentation\_block:

pyruvate kinase, plastid - common tobacco

C:Species: *Nicotiana tabacum* (common tobacco)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999

C:Accession: S44287

R:Blakeley, S.; Gottlob-McHugh, S.; Wan, J.; Crews, L.; Mikl, B.; Ko, K.; Dennis, D.

submitted to the EMBL Data Library, November 1993

A:Description: Molecular characterisation of plastid pyruvate kinase from castor and

A:Reference number: S44286

A:Accession: S44287

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-562 <BLA>

A:Cross-references: EMBL:Z28374; NID:g482937; PIDN:CAA82223.1; PID:g482938

C:Superfamily: pyruvate kinase

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S44287 ..

Align seg 1/1 to: S44287 from: 1 to: 562

627 AACTGTTATTATCGAATCCTTG 650

|||||

522 LysLeuLeuSerLysSerLeu 529

seq\_name: pir2:E96598

seq\_documentation\_block:

protein F20N2.2 [Imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: E96598

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hulzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719

A:Accession: E96598

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-607 <STO>

A:Cross-references: GB:AE005173; NID:g8778504; PIDN:AAF79512.1; GSPDB:GN00141

C:Genetics:

A:Gene: F20N2.2

A:Map position: 1

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x E96598 ..

Align seg 1/1 to: E96598 from: 1 to: 607

1276 AATCTTGAAGTTCTTGATCTTGGC 1299  
|||||  
389 AsnLeuGluValLeuAspLeuGly 396

seq\_name: p1r2:T10727

seq\_documentation\_block:  
protein kinase Xa21 (PC 2.7.1.1.-) D, receptor type - long-staminate rice  
C:Species: Oryza longistaminata (long-staminate rice)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T10727  
R:Song, W.Y.; Pi, L.Y.; Wang, G.L.; Gardner, J.; Holsten, T.; Ronald, P.C.  
Plant Cell 9, 1279-1287, 1997  
A:Title: Evolution of the rice Xa21 disease resistance gene family.  
A:Reference number: 215276; MUID:97432142  
A:Accession: T10727  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-612 <SON>  
A:Cross-references: EMBL:U72726; NID:g2586078; PIDN:AAB82753.1; PID:g2586081  
A:Experimental source: strain IRBB21  
C:Genetics:  
A:Map position: 11  
C:Keywords: phosphotransferase

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x T10727 ..

Align seg 1/1 to: T10727 from: 1 to: 612

2194 TTGAGCTCAGCCACCAACCACTG 2217  
|||||  
156 LeuAspLeuSerHisnGlnLeu 163

seq\_name: p1r2:S74727

seq\_documentation\_block:  
precorrin methylase (PC 2.1.1.1.-) - Synecchocystis sp. (strain PCC 6803)  
N:Alternate names: protein sir0569  
C:Species: Synecchocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997  
C:Accession: S74727  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S74727  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-627 <KAN>  
A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAAL6878.1; PID:d101761  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: cbhH  
C:Keywords: methyltransferase

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S74727 ..

Align seg 1/1 to: S74727 from: 1 to: 627

627 AACTCTTATTATCGAAATCCTTG 650  
|||||  
42 LysLeuLeuSerLysSerLeu 49

seq\_name: p1r2:S47299

seq\_documentation\_block:  
gene F protein - rinderpest virus  
C:Species: rinderpest virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Sep-1999  
C:Accession: S47299  
R:Evans, S.A.; Barton, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A:Description: The complete nucleotide sequence of the fusion protein gene of the vac  
A:Reference number: S47299  
A:Accession: S47299  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-636 <EVA>  
A:Cross-references: EMBL:Z31655; NID:g535391; PIDN:CAA83481.1; PID:g535392  
C:Superfamily: parainfluenza virus cell fusion protein

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S47299 ..

Align seg 1/1 to: S47299 from: 1 to: 636

2491 GTGACTATTCTTACCTGGCCACA 2514  
|||||  
101 ValThrIleProTyrLeuAlaThr 108

seq\_name: p1r2:T47895

seq\_documentation\_block:  
hypothetical protein T4C21.250 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47895  
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: 224479  
A:Accession: T47895  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-648 <CHO>  
A:Cross-references: EMBL:AL162295  
A:Experimental source: cultivar Columbia; BAC clone T4C21  
C:Genetics:

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x T47895 ..

Align seg 1/1 to: T47895 from: 1 to: 648

1937 GTGAGTCTCTTAGAACTCTGGAAT 1960  
|||||  
220 ValSerLeuLeuGluLeuTrpAsn 227

seq\_name: pir2:G89894

seq\_documentation\_block:  
protein kinase [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: G89894

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A99758; PMID:21311952; PMID:11418146

A:Accession: G89894

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-664 <KUR>

A:Cross-references: GB:BA000018; PID:gl3701020; PIDN:BA842315.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA1063

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x G89894 ..

Align seg 1/1 to: G89894 from: 1 to: 664

767 GTTAAAGTAGATGGCAAAACAGTA 744

|||||

647 ValIysValAspGlyIysThrVal 654

seq\_name: pir2:S04128

seq\_documentation\_block:

phenylalanine ammonia-lyase (EC 4.3.1.5) class III - kidney bean

C:Species: *Phaseolus vulgaris* (kidney bean)

C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 21-May-1999

C:Accession: S04128

R:Cramer, C.L.; Edwards, K.; Dron, M.; Liang, X.; Dildine, S.L.; Bolwell, G.P.; Dixon, R

Plant Mol. Biol. 12, 367-383, 1989

A:Title: Phenylalanine ammonia-lyase gene organization and structure.

A:Reference number: S04127

A:Accession: S04128

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-710 <CRA>

C:Genetics:

A:Gene: PAL3

C:Superfamily: histidine ammonia-lyase

C:Keywords: ammonia-lyase; carbon-nitrogen lyase

F:198-200/Cross-link: 5-Imidazolinone (Ala-Gly) #status predicted

F:199/Modified site: dehydroalanine (Ser) #status predicted

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x S04128 ..

Align seg 1/1 to: S04128 from: 1 to: 710

2144 TCAAAATCTTTCAGTTCGAAGAAC 2167

|||||

436 SerAsnLeuSerValGlyArgAsn 443

seq\_name: pir2:C84633

seq\_documentation\_block:

probable disease resistance protein [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: C84633

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487

A:Accession: C84633

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-743 <STO>

A:Cross-references: GB:AB002093; NID:g4115376; PIDN:AAD03377.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g24160

A:Map position: 2

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x C84633 ..

Align seg 1/1 to: C84633 from: 1 to: 743

2194 TTGGACCTCAGCCACCAACCACTG 2217

|||||

608 LeuAspLeuSerHisAsnGlnLeu 615







SwissProt_40:DROC_MYCPN +	7.00	100.09	314.84	224	P09924	mycoplasma pneumoniae	SwissProt_40:CHS4_BROFI -	7.00	96.16	236.24	394	O23730	bromheadia finlays
SwissProt_40:IRIL_FCMA -	7.00	99.78	313.36	234	P16809	human cytomegalovirus	SwissProt_40:CHS8_BROFI -	7.00	96.16	236.24	394	O23731	bromheadia finlays
SwissProt_40:YAIL_LACVA -	7.00	99.72	313.08	236	O9CJb2	lactococcus lactis (s	SwissProt_40:DRD2_BROFI -	7.00	96.14	236.16	395	P48988	drosophila melanog
SwissProt_40:YML1_YEAST +	7.00	99.69	312.93	237	P40207	saccharomyces cerevis	SwissProt_40:YK08_CAEEL +	7.00	96.11	236.00	397	P34303	caenorhabditis ele
SwissProt_40:YML7_YEAST +	7.00	99.64	312.65	239	P40113	arabidopsis thaliana	SwissProt_40:CAE3_CANAL +	7.00	96.09	235.92	398	P43092	candida albicans (
SwissProt_40:BIL_DROME +	7.00	99.46	311.82	245	O9Vsh3	drosophila melanogaste	SwissProt_40:PCNB_BUCAL +	7.00	96.07	235.84	399	P57442	buchnera aphidicol
SwissProt_40:YAO6_SCHPO +	7.00	99.35	311.27	249	O10106	schizosaccharomyces p	SwissProt_40:YXAM_BACSU +	7.00	96.07	235.84	399	P42112	bacillus subtilis.
SwissProt_40:DLX1_LACVA -	7.00	99.18	310.47	255	O64317	mus musculus (mouse).	SwissProt_40:HOFC_BACSU +	7.00	96.05	235.76	400	P36646	escherichia coli.
SwissProt_40:Y139_BUCAI -	7.00	99.16	310.34	256	P57239	buchnera aphidicola (	SwissProt_40:H136_ARATH +	7.00	96.00	235.52	403	O82660	arabidopsis thalia
SwissProt_40:YD69_AQUAE -	7.00	99.05	309.82	260	O67381	aquifex aeolicus (hyper	SwissProt_40:LM06_HUMAN +	7.00	95.93	235.21	407	O43900	homo sapiens (huma
SwissProt_40:LM45_HUMAN -	7.00	99.00	309.57	262	O13371	homo sapiens (human).	SwissProt_40:HID_DROME +	7.00	95.88	234.98	410	O24300	drosophila melanoga
SwissProt_40:CBEL_FINTH +	7.00	98.89	309.06	266	P10049	pinus thunbergii (gre	SwissProt_40:SNC3_HUMAN +	7.00	95.86	234.90	411	O92966	homo sapiens (huma
SwissProt_40:PERA_ECO27 +	7.00	98.68	308.08	274	P43459	escherichia coli ol27	SwissProt_40:AIHAF_RABIT -	7.00	95.83	234.74	413	P23035	oryctolagus cunicu
SwissProt_40:RPE_ORISA -	7.00	98.68	308.08	274	O92atp5	oryza sativa (rice).	SwissProt_40:CDSD_THEMEA	7.00	95.83	234.74	413	O9X191	thermotoga maritima
SwissProt_40:KUSA_RICPR -	7.00	98.66	307.96	275	O9ze84	rickettsia prowazekii	SwissProt_40:SYS_MYCLE -	7.00	95.76	234.44	417	O9cdcl	mycobacterium lepra
SwissProt_40:LECS_DOLBI -	7.00	98.66	307.96	275	P19588	dolichos biflorus (ho	SwissProt_40:SYS_MYCTU -	7.00	95.73	234.29	419	P66244	mycobacterium tuber
SwissProt_40:PHAE_PHAVU -	7.00	98.66	307.96	275	P05088	phaseolus vulgaris (M	SwissProt_40:DCDA_ECOLI +	7.00	95.71	234.21	420	P00861	escherichia coli.
SwissProt_40:PHAE_LACIC +	7.00	98.56	307.48	279	P43909	lactococcus lactis (s	SwissProt_40:OD02_BUCAL +	7.00	95.71	234.21	420	P57389	buchnera aphidicol
SwissProt_40:YICC_ECOLI +	7.00	98.36	306.54	287	P23839	escherichia coli. pro	SwissProt_40:ACDM_MOUSE -	7.00	95.70	234.13	421	P43592	mus musculus (mous
SwissProt_40:ARY1_RAT -	7.00	98.29	306.20	290	P50297	rattus norvegicus (rat)	SwissProt_40:ACDM_PIG -	7.00	95.70	234.13	421	P41367	sus scrofa (pig). ac
SwissProt_40:MT51_METVA +	7.00	97.83	304.00	310	O57983	mycoplasma jannasch	SwissProt_40:ACDM_RAT -	7.00	95.70	234.13	421	P08503	rattus norvegicus (r
SwissProt_40:Y302_MYCPN +	7.00	97.67	303.27	317	P75357	mycoplasma pneumoniae	SwissProt_40:MURA_THEMEA +	7.00	95.66	233.98	423	O9XW33	thermotoga maritim
SwissProt_40:KOGT_ERWCA -	7.00	97.65	303.17	318	O9XB52	erwinia carotovora. 2	SwissProt_40:STH_MYCTU +	7.00	95.66	233.98	423	O50641	mycobacterium tuber
SwissProt_40:NUIM_PHACI +	7.00	97.65	303.17	318	O78706	phascoglossus cinereu	SwissProt_40:IBHA_BOVIN -	7.00	95.63	233.83	425	P07995	bos taurus (bovine
SwissProt_40:Y692_METTH +	7.00	97.65	303.17	318	O26788	methanobacterium the	SwissProt_40:OYB_SHEEP -	7.00	95.63	233.83	425	P43032	ovis aries (sheep)
SwissProt_40:YDDK_ECOLI +	7.00	97.65	303.17	318	P76123	escherichia coli. hyp	SwissProt_40:SVH_MYCLE +	7.00	95.60	233.69	427	P46696	mycobacterium lepra
SwissProt_40:EDMS_STRPR +	7.00	97.63	303.07	319	P45439	streptomyces fradiae	SwissProt_40:MTAL_ACEPA +	7.00	95.57	233.54	429	O52702	acetobacter pasteu
SwissProt_40:KPR3_YEAST +	7.00	97.60	302.96	320	P38689	saccharomyces cerevis	SwissProt_40:SP54_THEAO +	7.00	95.57	233.54	429	O07347	thermus aquaticus.
SwissProt_40:CCCL_YEAST +	7.00	97.56	302.76	322	P47818	saccharomyces cerevis	SwissProt_40:SERC_ARATH +	7.00	95.55	233.46	430	O96255	arabidopsis thalia
SwissProt_40:YCCP_ECOLI +	7.00	97.54	302.66	323	P76291	escherichia coli. hyp	SwissProt_40:HXB3_HUMAN +	7.00	95.53	233.39	431	P14651	homo sapiens (huma
SwissProt_40:YFCK_BACSU -	7.00	97.50	302.46	325	P42422	bacillus subtilis. hy	SwissProt_40:MMT2_PSEA -	7.00	95.50	233.24	433	O9P026	mus musculus (mous
SwissProt_40:YFN_METJA +	7.00	97.48	302.36	326	O58839	methanococcus jannasch	SwissProt_40:MMT2_PSEA -	7.00	95.42	232.88	438	O9P026	pseudomonas aerugi
SwissProt_40:MDH_THEFL +	7.00	97.45	302.26	327	P10584	thermus aquaticus (sub	SwissProt_40:CYB_HAETN +	7.00	95.39	232.74	440	P44855	haemophilus influe
SwissProt_40:PARB_MYCPN +	7.00	97.33	301.67	333	O50201	mycobacterium leprae.	SwissProt_40:DCV8_DIODI +	7.00	95.36	232.59	442	P54639	dictyostellium disc
SwissProt_40:PIX2_CHICK +	7.00	97.33	301.67	333	O93385	gallus gallus (chicke	SwissProt_40:NU4M_DROME -	7.00	95.30	232.31	446	P18931	drosophila melanog
SwissProt_40:RYCA_METVA +	7.00	97.22	301.98	338	O60335	methanococcus jannasch	SwissProt_40:NU4M_DROYA +	7.00	95.30	232.31	446	P07707	drosophila yakuba
SwissProt_40:RYCA_MAIZE +	7.00	97.18	300.99	340	P20024	zea mays (maize).	SwissProt_40:DCDA_MYCTU +	7.00	95.28	232.24	447	P31848	mycobacterium tube
SwissProt_40:KTHA_ARATH -	7.00	97.18	300.99	340	O91X33	arabidopsis thaliana	SwissProt_40:COAB_STRPN +	7.00	95.25	232.10	449	P36498	streptococcus pneu
SwissProt_40:KPTA_VZVD +	7.00	97.16	300.90	341	P09250	varicella-zoster virus	SwissProt_40:GAAG_BOVIN -	7.00	95.22	231.96	451	P10063	bos taurus (bovine
SwissProt_40:T4BB_BACCO +	7.00	97.16	300.90	341	O76506	bacillus coagulans. 1	SwissProt_40:GAAG_HUMAN -	7.00	95.22	231.96	451	P47869	homo sapiens (huma
SwissProt_40:ARGC_BACCO +	7.00	97.06	300.52	346	O9K8V2	bacillus halodurans.	SwissProt_40:GAAG_MOUSE -	7.00	95.22	231.96	451	P26048	mus musculus (mous
SwissProt_40:MEB4_HUMAN +	7.00	97.06	300.42	346	Q15481	homo sapiens (human).	SwissProt_40:GAAG_RAT -	7.00	95.22	231.96	451	P23576	rattus norvegicus (r
SwissProt_40:YB09_YEAST +	7.00	97.04	300.33	347	P38286	saccharomyces cerevis	SwissProt_40:ACHS_RAT -	7.00	95.20	231.89	452	P20420	rattus norvegicus (r
SwissProt_40:KERA_BOVIN -	7.00	96.94	299.87	352	O62702	bos taurus (bovine).	SwissProt_40:CARA_NEUCR -	7.00	95.19	231.82	453	P22572	neurospora crassa.
SwissProt_40:KERA_HUMAN +	7.00	96.94	299.87	352	O60938	homo sapiens (human).	SwissProt_40:GAAL_CHICK -	7.00	95.16	231.68	455	P19150	gallus gallus (chi
SwissProt_40:KERA_CHICK +	7.00	96.92	299.78	353	O42235	gallus gallus (chicke	SwissProt_40:GAAL_MOUSE -	7.00	95.16	231.68	455	P18504	mus musculus (mous
SwissProt_40:KERA_COTJA +	7.00	96.92	299.78	353	O9de66	coturnix coturnix (ja	SwissProt_40:NH55_CAEEL +	7.00	95.16	231.68	455	O16962	caenorhabditis ele
SwissProt_40:Y120_RICPR +	7.00	96.90	299.68	354	O9ze28	rickettsia prowazekii	SwissProt_40:Y955_MYCTU -	7.00	95.16	231.68	455	P71555	mycobacterium tube
SwissProt_40:ALFB_CHICK +	7.00	96.73	298.87	363	P07341	gallus gallus (chicke	SwissProt_40:GAAL_BOVIN -	7.00	95.14	231.61	456	P08219	bos taurus (bovine
SwissProt_40:Y632_METUA -	7.00	96.73	298.87	363	O58049	methanococcus jannasch	SwissProt_40:GAAL_HUMAN -	7.00	95.14	231.61	456	P14867	homo sapiens (huma
SwissProt_40:HRCA_XYLF -	7.00	96.65	298.52	367	O9pb03	xyella fastidiosa. H	SwissProt_40:YCL3_ASTIO +	7.00	95.14	231.61	456	P14761	astasia longa (eug
SwissProt_40:Y093_HAETN -	7.00	96.63	298.43	368	P44509	haemophilus influenza	SwissProt_40:ARLY_MYCTU -	7.00	95.13	231.54	457	O52733	lactobacillus brev
SwissProt_40:WNT1_AMBME +	7.00	96.61	298.35	369	P21551	ambystoma mexicanum	SwissProt_40:PPB4_BACSU +	7.00	95.07	231.27	461	P19406	bacillus subtilis.
SwissProt_40:WNT1_HUMAN +	7.00	96.60	298.26	370	P04628	homo sapiens (human).	SwissProt_40:BLND_LYIYA +	7.00	95.05	231.20	462	P23117	lytechinus variega
SwissProt_40:WNT1_MOUSE +	7.00	96.60	298.26	370	P04426	mus musculus (mouse).	SwissProt_40:PUSB_CARTI +	7.00	95.04	231.13	463	O42713	carthamus tinctori
SwissProt_40:WNT1_XENIA +	7.00	96.58	298.17	371	P10108	xenopus laevis (afric	SwissProt_40:UHPY_ECOLI +	7.00	95.04	231.13	463	P13408	escherichia coli.
SwissProt_40:YPC4_CAEEL +	7.00	96.58	298.17	371	Q11181	caenorhabditis elegans	SwissProt_40:ARLY_MYCTU -	7.00	94.93	230.66	470	P94934	mycobacterium tube
SwissProt_40:ACT_GIALA +	7.00	96.50	297.83	375	P51775	glardia lambia (glard	SwissProt_40:DCDA_MYCLE -	7.00	94.90	230.53	472	O50140	mycobacterium lepr
SwissProt_40:KLAB_ECOLI -	7.00	96.45	297.57	378	O52328	escherichia coli. kla	SwissProt_40:PCCB_MYCTU +	7.00	94.89	230.46	473	O10506	mycobacterium tube
SwissProt_40:SYMM_YEAST +	7.00	96.43	297.49	379	O4803	saccharomyces cerevis	SwissProt_40:DCDA_MYCYM +	7.00	94.87	230.40	473	Q9X5M1	mycobacterium smeg
SwissProt_40:Y2B8_SYNT3 -	7.00	96.43	297.49	379	Q35518	synthocystis sp. (st	SwissProt_40:UL49_HSVJ +	7.00	94.87	230.20	477	P52442	human herpesvirus
SwissProt_40:CYB_CARPL +	7.00	96.39	297.32	381	P34866	carcinarius porus (s	SwissProt_40:UL49_HSVJ +	7.00	94.87	230.20	477	P52442	human herpesvirus
SwissProt_40:CYB_CARPO +	7.00	96.39	297.32	381	P34867	carcinarius porus (s	SwissProt_40:MM20_BOVIN -	7.00	94.71	289.94	481	O18767	bos taurus (bovine
SwissProt_40:CYB_GALCU +	7.00	96.39	297.32	381	P34868	galiecerdo cuvier (tig	SwissProt_40:Y136_TREPA +	7.00	94.71	289.68	485	O83172	treponema pallidum
SwissProt_40:CYB_NEGRB +	7.00	96.39	297.32	381	P34872	neqarion brevisotris	SwissProt_40:COGL_HUMAN +	7.00	94.70	289.62	486	O92879	homo sapiens (huma
SwissProt_40:CYB_PRIGL +	7.00	96.39	297.32	381	P34873	prionace glauca (blue	SwissProt_40:CUGL_MOUSE +	7.00	94.70	289.62	486	P28659	mus musculus (mous
SwissProt_40:CYB_SPHLE +	7.00	96.39	297.32	381	P34874	sphyrna lewini (hammer	SwissProt_40:LACL_BOYCI -	7.00	94.70	289.62	486	Q12570	botrytis cinerea (
SwissProt_40:CYB_SPHHT +	7.00	96.39	297.32	381	P34874	sphyrna tiburo (tiburo	SwissProt_40:MDM4_HUMAN +	7.00	94.64	289.36	490	O15151	homo sapiens (huma
SwissProt_40:CDAR_ECOLI -	7.00	96.32	296.98	385	P34875	sphyrna tiburo vesper	SwissProt_40:VNUC_IADHK +	7.00	94.61	289.23	492	O92937	chlamydia pneumoni
SwissProt_40:SASA_SYNT3 +	7.00	96.28	296.82	387	P37047	escherichia coli. car	SwissProt_40:VNUC_IADHK +	7.00	94.53	288.86	498	P16978	influenza a virus
SwissProt_40:YMR2_YEAST +	7.00	96.25	296.65	389	O06904	saccharomyces sp. (st	SwissProt_40:PROP_ECOLI -	7.00	94.50	288.73	500	O03017	drosophila melanog
SwissProt_40:YMR3_YEAST +	7.00	96.25	296.65	389	O04951	saccharomyces cerevis	SwissProt_40:VNUC_IADHK +	7.00	94.50	288.73	500	P30848	escherichia coli.
SwissProt_40:CHS3_BROFI -	7.00	96.16	236.24	394	O23729	bromheadia finlaysonii	SwissProt_40:VNN3_HUMAN +	7.00	94.50	288.67	501	P40862	salmonella typhimu



SwissProt_40:UAP1_ARATH -	7.00	94.47	288.61	502	! 064765 arabidopsis thaliana	SwissProt_40:HUTH_HUMAN +	7.00	92.60	280.35	657	! P42357 homo sapiens (huma
SwissProt_40:YH60_MYCTU +	7.00	94.47	288.61	502	! 006795 mycobacterium tubercu	SwissProt_40:PGSL_SALTY +	7.00	92.49	279.85	668	! P37433 salmonella typhimu
SwissProt_40:FLIC_SALBU +	7.00	94.45	288.48	504	! 006969 salmonella budapestc	SwissProt_40:PTB1_YEAST +	7.00	92.48	279.81	669	! P23900 saccharomyces cere
SwissProt_40:FLIC_SALDE +	7.00	94.45	288.48	504	! 006970 salmonella derby fla	SwissProt_40:LOX5_MESAU -	7.00	92.44	279.67	672	! P51399 mesocricetus aurat
SwissProt_40:FLIC_SALDU +	7.00	94.45	288.48	504	! 006971 salmonella dublin fla	SwissProt_40:LOX5_RAT -	7.00	92.44	279.67	672	! P12527 rattus norvegicus (r
SwissProt_40:FLIC_SALDN +	7.00	94.45	288.48	504	! 006972 salmonella enteritidis	SwissProt_40:LOX5_MOUSE -	7.00	92.43	279.63	673	! P48999 mus musculus (mous
SwissProt_40:FLIC_SALMC +	7.00	94.45	288.48	504	! 006981 salmonella moscow fla	SwissProt_40:VTER_HCMVA +	7.00	92.42	279.58	674	! P16732 human cytomegalovi
SwissProt_40:FLIC_SALMO +	7.00	94.45	288.48	504	! 006973 salmonella montevideo	SwissProt_40:H57C_TRYVB +	7.00	92.40	279.49	676	! P20030 trypanosoma brucei
SwissProt_40:FLIC_SALNA +	7.00	94.45	288.48	504	! 052959 salmonella naestved	SwissProt_40:YD64_MYCPN +	7.00	92.39	279.45	677	! P75417 mycoplasma pneumon
SwissProt_40:FLIC_SALRO +	7.00	94.45	288.48	504	! 006982 salmonella rostock fla	SwissProt_40:HKR3_HUMAN +	7.00	92.28	278.95	688	! P10074 homo sapiens (huma
SwissProt_40:FLIC_SALSE +	7.00	94.45	288.48	504	! 006983 salmonella senftenber	SwissProt_40:VAT1_DEIRA -	7.00	92.26	278.87	690	! Q9RWH3 deinococcus radiod
SwissProt_40:ICL1_MOUSE +	7.00	94.45	288.48	504	! P97290 mus musculus (mouse)	SwissProt_40:YG29_YEAST -	7.00	92.12	278.27	704	! P36061 saccharomyces cere
SwissProt_40:GAE1_RAT +	7.00	94.42	288.36	506	! Q9ES14 rattus norvegicus (rat)	SwissProt_40:YK70_YEAST -	7.00	92.12	278.19	706	! P36166 saccharomyces cere
SwissProt_40:FLIC_SALBE +	7.00	94.40	288.30	507	! 006978 salmonella bertea fla	SwissProt_40:GAC1_HUMAN +	7.00	92.08	278.10	708	! P57273 buchnera aphidicola
SwissProt_40:FLIC_SALON +	7.00	94.40	288.30	507	! 006974 salmonella oranienber	SwissProt_40:GAC1_HUMAN +	7.00	92.03	277.89	713	! P75735 homo sapiens (huma
SwissProt_40:GSH1_MEDFR +	7.00	94.39	288.24	508	! Q9ZNX6 medicago truncatula c	SwissProt_40:MRPL_ARATH +	7.00	91.90	277.31	727	! Q9LW85 arabidopsis thalia
SwissProt_40:MS2P_CRIGR +	7.00	94.36	288.11	510	! 054862 cricetus griseus (c	SwissProt_40:ACPH_HUMAN +	7.00	91.85	277.10	732	! P13798 homo sapiens (huma
SwissProt_40:DOC1_HAEIN +	7.00	94.35	288.05	511	! P71362 haemophilus influenzae	SwissProt_40:ACPH_PIG +	7.00	91.85	277.10	732	! P19205 sus scrofa (pig)
SwissProt_40:ANM3_HUMAN +	7.00	94.34	287.99	512	! 050678 homo sapiens (human)	SwissProt_40:ACPH_RAT +	7.00	91.85	277.10	732	! P13676 rattus norvegicus (r
SwissProt_40:CGD2_DROME +	7.00	94.34	287.99	512	! Q9W223 drosophila melanogast	SwissProt_40:ADDI_CAEEL +	7.00	91.85	277.10	732	! Q9U9K0 caenorhabditis ele
SwissProt_40:CP11_HUMAN +	7.00	94.34	287.99	512	! P04798 homo sapiens (human)	SwissProt_40:YJ32_YEAST -	7.00	91.84	277.06	733	! P43602 saccharomyces cere
SwissProt_40:CP11_MACFA +	7.00	94.34	287.99	512	! P33616 macaca fascicularis c	SwissProt_40:GLJ2_DICDI +	7.00	91.83	277.02	734	! P34116 dictyostelium disc
SwissProt_40:HMW2_ARATH +	7.00	94.34	287.99	512	! Q04921 arabidopsis thaliana	SwissProt_40:TRFM_HUMAN +	7.00	91.79	276.86	738	! P08582 homo sapiens (huma
SwissProt_40:LAT1_MOUSE +	7.00	94.34	287.99	512	! Q9Z127 mus musculus (mouse)	SwissProt_40:KHL1_HUMAN +	7.00	91.70	276.46	748	! Q9NR54 homo sapiens (huma
SwissProt_40:FLU2_ARATH +	7.00	94.32	287.93	513	! Q9FX97 arabidopsis thaliana	SwissProt_40:TALA_POVHA -	7.00	91.67	276.34	751	! P03075 hamster polyomavir
SwissProt_40:SNX2_HUMAN +	7.00	94.24	287.57	519	! 060749 homo sapiens (human)	SwissProt_40:PXAL_YEAST -	7.00	91.61	276.06	758	! P41909 saccharomyces cere
SwissProt_40:SNX2_MOUSE +	7.00	94.24	287.57	519	! Q9CWK8 mus musculus (mouse)	SwissProt_40:ORPA_HUMAN +	7.00	91.55	275.83	764	! Q9BXB5 homo sapiens (huma
SwissProt_40:SYN1_TREPA +	7.00	94.19	287.33	523	! 083618 treponema pallidum, as	SwissProt_40:LEM3_RAT -	7.00	91.52	275.67	768	! P98106 rattus norvegicus (r
SwissProt_40:ANM3_RAT +	7.00	94.12	287.04	528	! 070467 rattus norvegicus (rat)	SwissProt_40:COMP_BACSU -	7.00	91.51	275.63	769	! Q90927 bacillus subtilis
SwissProt_40:LAD1_MOUSE +	7.00	94.12	287.04	528	! P37016 mus musculus (mouse)	SwissProt_40:YXK0_YEAST +	7.00	91.49	275.56	771	! Q02208 saccharomyces cere
SwissProt_40:HLV1_HALI7 -	7.00	94.10	286.92	530	! P21913 halophilic bacteria st	SwissProt_40:RRL1_HSVBE -	7.00	91.32	274.83	790	! P28846 equine herpesvirus
SwissProt_40:TACY_LISSE +	7.00	94.10	286.92	530	! P31830 listeria seigelieri, s	SwissProt_40:ADDI_CAEEL +	7.00	91.29	274.72	793	! Q93917 dictyostelium disc
SwissProt_40:YD3D_SCHPO +	7.00	94.06	286.69	533	! Q10277 schizosaccharomyces p	SwissProt_40:YJ32_YEAST -	7.00	91.28	274.72	793	! Q93917 dictyostelium disc
SwissProt_40:CN1B_BOVIN +	7.00	94.04	286.69	534	! Q10161 bos taurus (bovine)	SwissProt_40:FGR3_HUMAN +	7.00	91.28	274.24	806	! P61851 mus musculus (mous
SwissProt_40:SP70_DICDI +	7.00	94.00	286.52	537	! P15269 dictyostelium discoide	SwissProt_40:TLRA_HUMAN +	7.00	91.14	274.06	811	! Q9BXR5 homo sapiens (huma
SwissProt_40:YEJL_ECOLI -	7.00	93.99	286.46	538	! P52127 escherichia coli, hyp	SwissProt_40:SWI3_YEAST -	7.00	91.02	273.55	825	! P32591 saccharomyces cere
SwissProt_40:HNAL_ARATH +	7.00	93.93	286.17	543	! P42804 arabidopsis thaliana	SwissProt_40:SIPI4_YEAST +	7.00	90.98	273.41	829	! P46954 saccharomyces cere
SwissProt_40:LEM2_RAT -	7.00	93.85	285.83	549	! P98105 rattus norvegicus (rat)	SwissProt_40:YJG2_YEAST +	7.00	90.97	273.37	830	! P40367 rattus norvegicus (r
SwissProt_40:PQO1_YEAST +	7.00	93.85	285.83	549	! P32945 saccharomyces cerevis	SwissProt_40:TLR4_RAT +	7.00	90.93	273.20	835	! Q9QX05 rattus norvegicus (r
SwissProt_40:TPTE_HUMAN +	7.00	93.82	285.72	551	! P56180 homo sapiens (human)	SwissProt_40:TLR4_CRIGR +	7.00	90.91	273.09	838	! Q9W882 cricetus griseus
SwissProt_40:YGF1_YEAST +	7.00	93.82	285.72	551	! P53214 saccharomyces cerevis	SwissProt_40:PULA_THEMA -	7.00	90.87	272.92	843	! Q93840 thermotoga maritim
SwissProt_40:Y423_MYCCE +	7.00	93.70	285.17	561	! P47662 mycoplasma genitalium	SwissProt_40:AMPN_LACHE -	7.00	90.86	272.88	844	! Q10730 lactobacillus helv
SwissProt_40:SYR1_CHLMU -	7.00	93.68	285.06	563	! Q9PJ18 chlamydia muridarum, a	SwissProt_40:ENV_HV2SB +	7.00	90.84	272.81	846	! P12449 human immunodefici
SwissProt_40:SYR1_CHLTR -	7.00	93.68	285.06	563	! 084466 chlamydia trachomatis	SwissProt_40:ENV_HV2D1 +	7.00	90.80	272.64	851	! P17755 human immunodefici
SwissProt_40:ARAB_BACST +	7.00	93.66	285.00	564	! Q9S468 bacillus stearothermo	SwissProt_40:GAF1_SCHPO -	7.00	90.77	272.50	855	! Q10280 schizosaccharomyce
SwissProt_40:URBN_BACST +	7.00	93.65	284.95	565	! Q9UK80 homo sapiens (human)	SwissProt_40:ENV_HV2KR +	7.00	90.75	272.43	857	! Q74126 human immunodefici
SwissProt_40:HEMA_IAEN7 +	7.00	93.64	284.90	566	! P03440 influenza a virus (st	SwissProt_40:PMS2_MOUSE +	7.00	90.74	272.36	859	! P54279 mus musculus (mous
SwissProt_40:URBN_MOUSE +	7.00	93.64	284.90	566	! Q9QZ16 mus musculus (mouse)	SwissProt_40:ENV_HV2BE +	7.00	90.74	272.36	859	! Q9J1F7 mus musculus (mous
SwissProt_40:DPOL_BPCPI +	7.00	93.61	284.79	568	! Q37989 bacteriophage cp-1, o	SwissProt_40:ENV_HV2BE +	7.00	90.73	272.33	860	! P18094 human immunodefici
SwissProt_40:ESR2_ICTPU +	7.00	93.53	284.41	575	! Q9IAK1 ictalurus punctatus (	SwissProt_40:AMPN_CAUCR -	7.00	90.70	272.23	863	! P37893 caulobacter cresce
SwissProt_40:SHO2_HUMAN +	7.00	93.44	284.04	582	! Q9UQ13 homo sapiens (human)	SwissProt_40:PEM1_YEAST -	7.00	90.66	272.02	869	! P05374 saccharomyces cere
SwissProt_40:SHO2_MOUSE +	7.00	93.44	284.04	582	! Q88520 mus musculus (mouse)	SwissProt_40:LDVR_MOUSE +	7.00	90.62	271.89	873	! P98156 mus musculus (mous
SwissProt_40:FD26_MYCTU -	7.00	93.43	283.99	583	! Q10976 mycobacterium tubercu	SwissProt_40:LDVR_RABIT +	7.00	90.62	271.89	873	! P33593 oryctolagus cunicu
SwissProt_40:AYYL_MOUSE -	7.00	93.37	283.73	588	! Q9CTG6 mus musculus (mouse)	SwissProt_40:LDVR_RAT +	7.00	90.62	271.89	873	! P98166 rattus norvegicus (r
SwissProt_40:YMW7_YEAST +	7.00	93.35	283.62	590	! P33863 saccharomyces cerevis	SwissProt_40:YPD8_CAEEL +	7.00	90.62	271.89	873	! Q11187 caenorhabditis ele
SwissProt_40:TDPH_CAEEL +	7.00	93.26	283.21	598	! Q22703 caenorhabditis elegans	SwissProt_40:MUTS_SYNY3 -	7.00	90.58	271.72	878	! P73769 synecocystis sp.
SwissProt_40:TDPH_CAEEL +	7.00	93.26	283.21	598	! Q22703 caenorhabditis elegans	SwissProt_40:LYTD_BACSU +	7.00	90.57	271.66	880	! P39848 bacillus subtilis
SwissProt_40:HMW3_MYCCE +	7.00	93.24	283.16	599	! Q57081 mycoplasma genitalium	SwissProt_40:YDGH_BACSU +	7.00	90.53	271.49	885	! P96706 bacillus subtilis
SwissProt_40:SYD1_SYNY3 -	7.00	93.24	283.16	599	! P73851 synecocystis sp. (str	SwissProt_40:YB33_SCHPO +	7.00	90.48	271.29	891	! Q14338 schizosaccharomyce
SwissProt_40:SR68_DROME -	7.00	93.19	282.91	604	! Q9V392 drosophila melanogast	SwissProt_40:SECA_CYACA -	7.00	90.48	271.16	895	! Q14911 cyanidium caldarii
SwissProt_40:NRD1_HUMAN +	7.00	93.07	282.41	614	! P20393 homo sapiens (human)	SwissProt_40:EP15_HUMAN -	7.00	90.44	271.13	896	! P42566 homo sapiens (huma
SwissProt_40:CTRL1_CITPR -	7.00	93.03	282.21	618	! P23182 citrobacter freundii	SwissProt_40:RLR3_HUMAN +	7.00	90.42	271.03	899	! P33568 rattus norvegicus (ra
SwissProt_40:PDMA_CAEEL -	7.00	93.03	282.21	618	! P34329 caenorhabditis elegans	SwissProt_40:RLR3_HUMAN +	7.00	90.38	270.87	904	! Q15455 homo sapiens (huma
SwissProt_40:MTN4_HUMAN -	7.00	93.02	282.16	619	! Q95460 homo sapiens (human)	SwissProt_40:CO4_BOVIN -	7.00	90.26	270.36	920	! P01030 bos taurus (bovine)
SwissProt_40:FTSH_GUTH +	7.00	92.88	281.58	631	! Q78516 guillardia theta (cry	SwissProt_40:RB_MOUSE -	7.00	90.25	270.33	921	! P13405 mus musculus (mouse)
SwissProt_40:IL16_HUMAN +	7.00	92.88	281.58	631	! Q14005 homo sapiens (human)	SwissProt_40:RB_MOUSE -	7.00	90.20	270.10	928	! P06400 homo sapiens (human)
SwissProt_40:HTPG_PSEAE +	7.00	92.85	281.43	634	! Q91365 pseudomonas aeruginos	SwissProt_40:EAE_ECO57 -	7.00	90.15	269.92	934	! P43261 escherichia coli o1
SwissProt_40:SET1_MOUSE +	7.00	92.77	281.10	641	! Q9CS74 mus musculus (mouse)	SwissProt_40:EAE_ECO11 -	7.00	90.15	269.89	935	! Q31000 escherichia coli o1
SwissProt_40:R640_CAEEL -	7.00	92.75	281.00	643	! Q27274 caenorhabditis elegans	SwissProt_40:VP2_BTV11 -	7.00	89.99	269.24	956	! P05308 bluetongue virus (s
SwissProt_40:VP40_HSV2 +	7.00	92.75	281.00	643	! P52369 equine herpesvirus ty	SwissProt_40:SV4_HUMAN -	7.00	89.91	268.88	968	! P49588 homo sapiens (huma
SwissProt_40:DNM3_CHLPN +	7.00	92.74	280.96	644	! Q9Z6J9 chlamydia pneumoniae	SwissProt_40:MGR_DROME -	7.00	89.85	268.64	976	! P91685 drosophila melanoga
SwissProt_40:KNRL_DROME -	7.00	92.71	280.82	647	! P13054 drosophila melanogast	SwissProt_40:HIPI_HUMAN -	7.00	89.71	268.08	995	! Q00291 homo sapiens (huma
SwissProt_40:PGI1_YEAST +	7.00	92.70	280.77	648	! Q06216 saccharomyces cerevis	SwissProt_40:FRO_SORFI +	7.00	89.70	268.02	997	! Q09033 sordaria fimicola

Swissprot_40: CARE_MOUSE -	7.00	89.69	267.97	999	I Q99kfo mus musculus (mouse) .	Swissprot_40: HRX_HUMAN +	7.00	80.09	229.57	3969	I Q03164 homo sapiens, (huma
Swissprot_40: RLK5_ARATH +	7.00	89.51	267.97	999	I P47335 arabidopsis thaliana	Swissprot_40: DYHC_FUSO +	7.00	79.45	227.20	4349	I P78716 fusarium solani (
Swissprot_40: ADAL_YEAST +	7.00	89.51	267.97	1025	I P38065 saccharomyces cerevi	Swissprot_40: DYHC_NEUCR +	7.00	79.42	227.09	4367	I P45443 neurospora crassa
Swissprot_40: TLR9_MOUSE +	7.00	89.46	267.03	1032	I Q9eq3 mus musculus (mouse)	Swissprot_40: PGBM_HUMAN +	7.00	79.38	226.94	4393	I P98160 homo sapiens (hum
Swissprot_40: POLY_DROME +	7.00	89.44	266.94	1035	I P10401 drosophila melanogas	Swissprot_40: RRP4_CVMJH +	7.00	79.23	226.39	4488	I P19751 murine coronaviru
Swissprot_40: YKD3_YEAST +	7.00	89.42	266.86	1038	I P36097 saccharomyces cerevi	Swissprot_40: RYR3_HUMAN +	7.00	78.67	224.29	4870	I Q15413 homo sapiens (hum
Swissprot_40: HMD1_YEAST -	7.00	89.31	266.42	1051	I P12683 saccharomyces cerevi	Swissprot_40: CHEP_PARID +	6.00	101.22	4.7e+03	13	I P42718 parapolybia indica
Swissprot_40: ANPA_HUMAN -	7.00	89.27	266.23	1061	I P16066 homo sapiens (human)	Swissprot_40: ESTJ_MANSE +	6.00	100.22	4.6e+03	15	I P19985 manduca sexta (toba
Swissprot_40: TRC4_ECOLI -	7.00	89.27	266.23	1061	I P27189 escherichia coli, dr	Swissprot_40: RL11_STRAU -	6.00	100.22	4.6e+03	15	I Q9520 streptomyces aureof
Swissprot_40: TRC3A_STAAU -	7.00	89.24	266.12	1065	I Q9f1k0 staphylococcus aure	Swissprot_40: RL11_STRAU -	6.00	96.67	4.4e+03	25	I P82401 litoria raniformis
Swissprot_40: YCF3A_WARPO -	7.00	89.22	266.04	1068	I P12221 marchantia polymorph	Swissprot_40: AU52_LITRA -	6.00	96.67	4.4e+03	25	I P82402 litoria raniformis
Swissprot_40: HLES_DROME +	7.00	89.16	265.80	1077	I Q02308 drosophila melanogas	Swissprot_40: G3P2_JACOR +	6.00	96.67	4.4e+03	25	I P80447 jaculus orientalis
Swissprot_40: V120_HSV6U +	7.00	89.13	265.67	1082	I P52337 human herpesvirus (t	Swissprot_40: GBX1_MOUSE +	6.00	96.67	4.4e+03	25	I P82976 mus musculus (mouse
Swissprot_40: DPOL_ADE04 +	7.00	88.88	264.63	1122	I P05664 human adenovirus typ	Swissprot_40: C7L2_LITCI +	6.00	96.40	4.2e+03	26	I P81847 litoria citropa (hor
Swissprot_40: ALA3_ARATH +	7.00	88.87	264.61	1123	I Q9xie6 arabidopsis thaliana	Swissprot_40: C7L2_HORSE +	6.00	94.74	4.2e+03	33	I P81710 equus caballus (hor
Swissprot_40: LONK_YEAST +	7.00	88.81	264.35	1133	I P36775 saccharomyces cerevi	Swissprot_40: DP0B_BOVIN +	6.00	93.76	4.2e+03	38	I P27958 bos taurus (bovine)
Swissprot_40: KPCL_TREIR +	7.00	88.77	264.20	1139	I Q99014 trichoderma reesei (	Swissprot_40: H5_COLL1 +	6.00	93.76	4.2e+03	38	I P02260 columba livia (domest
Swissprot_40: MGCL_HUMAN +	7.00	88.76	264.13	1142	I Q60732 homo sapiens (human)	Swissprot_40: TYBB_ONCMY -	6.00	93.06	4.1e+03	42	I P26352 oncorhynchus mykiss
Swissprot_40: MMIC_MYCTU -	7.00	88.73	264.03	1146	I Q50585 mycobacterium tuber	Swissprot_40: YG71_METJA -	6.00	92.58	4.1e+03	45	I Q59065 methanococcus janna
Swissprot_40: YHC3_YEAST +	7.00	88.73	264.03	1146	I P38742 saccharomyces cerevi	Swissprot_40: VIT_ANAPL +	6.00	92.43	4.1e+03	46	I P56530 anas platyrhynchos
Swissprot_40: CD45_MOUSE +	7.00	88.69	263.88	1152	I P06800 mus musculus (mouse)	Swissprot_40: ATP8_PODAN +	6.00	91.85	4.1e+03	50	I Q02653 podospira anserina
Swissprot_40: PEX1_PICPA +	7.00	88.66	263.76	1157	I P46463 pichia pastoris (yea	Swissprot_40: LG2_TETTH -	6.00	91.71	4.0e+03	51	I P12072 tetrahymena thermoph
Swissprot_40: XPG_MOUSE -	7.00	88.59	263.44	1170	I P35689 mus musculus (mouse)	Swissprot_40: APD8_MYCSM +	6.00	91.58	4.0e+03	52	I Q50441 mycobacterium smegm
Swissprot_40: XPG_HUMAN -	7.00	88.49	263.05	1186	I P28715 homo sapiens (human)	Swissprot_40: SRY_AKOAZ +	6.00	91.58	4.0e+03	52	I P36388 akodon azarae. sex-d
Swissprot_40: SCIL_CHICK +	7.00	88.47	262.98	1189	I Q90988 gallus gallus (chick)	Swissprot_40: VQ66_BPHR2 -	6.00	91.31	4.0e+03	54	I P16516 bacteriophage phi-2
Swissprot_40: DPOL_ADE04 +	7.00	88.45	262.89	1193	I P87503 human adenovirus typ	Swissprot_40: VQ66_BPHR5 -	6.00	91.31	4.0e+03	54	I P15854 bacteriophage phi-1
Swissprot_40: XPG_XENLA -	7.00	88.43	262.82	1196	I P14629 xenopus laevis (afri	Swissprot_40: YP18_CLOPE -	6.00	91.19	4.0e+03	55	I P18018 clostridium perfrin
Swissprot_40: DP3A_CAMJE -	7.00	88.41	262.72	1200	I Q9p09 campylobacter jejuni	Swissprot_40: RL35_MYCGE +	6.00	90.70	4.0e+03	59	I P47339 escherichia genitai
Swissprot_40: XCFE_XENLA +	7.00	88.39	262.65	1203	I P50533 xenopus laevis (afri	Swissprot_40: YMP8_ECOLI +	6.00	90.58	4.0e+03	60	I P75979 escherichia coli, h
Swissprot_40: THR_DROME +	7.00	88.36	262.51	1209	I P42286 drosophila melanogast	Swissprot_40: CXIL_NAHJA -	6.00	90.35	4.0e+03	62	I P01472 najia haje annulifer
Swissprot_40: E75A_DROME +	7.00	88.20	261.86	1237	I P16771 drosophila melanogas	Swissprot_40: YC33_SKECO +	6.00	90.13	4.0e+03	64	I Q96808 skeletohemata costatu
Swissprot_40: FLI1_DROME +	7.00	88.09	261.43	1256	I Q24020 drosophila melanogas	Swissprot_40: RL30_MYCTU +	6.00	90.02	3.9e+03	65	I P595070 mycobacterium tuber
Swissprot_40: PER2_MOUSE -	7.00	88.09	261.41	1257	I Q54943 mus musculus (mouse)	Swissprot_40: YX27_PYRHO -	6.00	90.02	3.9e+03	65	I P78078 pyrococcus horikosh
Swissprot_40: S2AB_HUMAN +	7.00	88.03	261.16	1268	I Q95487 homo sapiens (human)	Swissprot_40: YC22_GUTH -	6.00	90.02	3.9e+03	65	I Q78421 guillardia theta (c
Swissprot_40: ICP4_HSV11 -	7.00	87.86	260.51	1298	I P08392 herpes simplex virus	Swissprot_40: THIS_ECOLI -	6.00	89.92	3.9e+03	66	I Q3583 escherichia coli, t
Swissprot_40: SWIL_YEAST +	7.00	87.78	260.16	1314	I P09547 saccharomyces cerevi	Swissprot_40: ATP8_ORYAF +	6.00	89.81	3.9e+03	67	I Q9xm19 oryctolopus afer (a
Swissprot_40: YDM5_SCHPO -	7.00	87.66	259.68	1337	I P87136 schizosaccharomyces	Swissprot_40: CMND_HAEIN -	6.00	89.81	3.9e+03	67	I P45035 haemophilus influen
Swissprot_40: Y014_CAEEL +	7.00	87.59	259.38	1351	I Q02331 caenorhabditis elega	Swissprot_40: Y002_BPHR1 +	6.00	89.71	3.9e+03	68	I P51701 bacteriophage hpl
Swissprot_40: M3K5_MOUSE -	7.00	87.47	258.91	1374	I Q99683 homo sapiens (human)	Swissprot_40: RL30_MYCLE +	6.00	89.41	3.9e+03	71	I Q33001 mycobacterium lepr
Swissprot_40: RON_MOUSE -	7.00	87.45	258.83	1378	I Q62190 mus musculus (mouse)	Swissprot_40: YF77_HAEIN +	6.00	89.31	3.9e+03	72	I Q57070 haemophilus influen
Swissprot_40: M3K5_MOUSE -	7.00	87.44	258.81	1379	I Q35099 mus musculus (mouse)	Swissprot_40: GYM2_HALN2 +	6.00	89.22	3.9e+03	73	I P33957 halobacterium sp. (
Swissprot_40: RPOB_ODOSI +	7.00	87.44	258.81	1379	I P49466 odontella sinensis .	Swissprot_40: RL17_BOVIN -	6.00	89.22	3.9e+03	73	I P82916 bos taurus (bovine)
Swissprot_40: N157_YEAST +	7.00	87.38	258.57	1391	I P40064 saccharomyces cerevi	Swissprot_40: WDMN_MOUSE -	6.00	89.12	3.9e+03	74	I Q62477 mus musculus (mouse
Swissprot_40: E75B_DROME +	7.00	87.37	258.51	1394	I P16772 drosophila melanogas	Swissprot_40: TRJ7_ECOLI +	6.00	89.03	3.9e+03	75	I P05835 escherichia coli, t
Swissprot_40: RON_HUMAN -	7.00	87.34	258.39	1400	I Q04912 homo sapiens (human)	Swissprot_40: UL91_HSVSA +	6.00	89.03	3.9e+03	75	I Q04193 herpesvirus saimiri
Swissprot_40: WC11_BOVIN -	7.00	87.16	257.68	1436	I P30205 bos taurus (bovine)	Swissprot_40: YF67_THEMA +	6.00	89.03	3.9e+03	75	I Q9xiq3 thermotoga maritima
Swissprot_40: E75C_DROME +	7.00	87.13	257.55	1443	I P13055 drosophila melanogas	Swissprot_40: I3EL_HUMAN +	6.00	88.94	3.9e+03	76	I Q43715 homo sapiens (human
Swissprot_40: BUD4_YEAST -	7.00	87.11	257.47	1447	I P47136 saccharomyces cerevi	Swissprot_40: CCEP_NITEU -	6.00	88.94	3.9e+03	76	I P59229 nitrosomonas europ
Swissprot_40: TRC5_ECOLI -	7.00	87.10	257.45	1448	I P27136 escherichia coli, dr	Swissprot_40: CSBA_BACSU -	6.00	88.94	3.9e+03	76	I P37953 bacillus subtilis .
Swissprot_40: KEM1_YEAST -	7.00	86.73	255.96	1528	I P22147 saccharomyces cerevi	Swissprot_40: TRJ7_ECOLI -	6.00	88.94	3.9e+03	76	I P33785 escherichia coli, t
Swissprot_40: BPT1_YEAST +	7.00	86.59	255.41	1559	I P14772 saccharomyces cerevi	Swissprot_40: DSRD_ARCFU +	6.00	88.85	3.9e+03	77	I P70742 archaeoglobus fulgi
Swissprot_40: YN12_YEAST +	7.00	86.44	254.83	1592	I P35855 saccharomyces cerevi	Swissprot_40: GP46_BPSPI +	6.00	88.85	3.9e+03	77	I Q48400 bacteriophage sp01
Swissprot_40: SNP2_YEAST -	7.00	85.98	252.99	1703	I P22082 saccharomyces cerevi	Swissprot_40: MINE_HELPY +	6.00	88.85	3.9e+03	77	I Q92ma7 helicobacter pylori
Swissprot_40: AKAC_HUMAN +	7.00	85.66	251.77	1781	I Q02952 homo sapiens (human)	Swissprot_40: MINE_HELPY +	6.00	88.85	3.9e+03	77	I Q25099 helicobacter pylori
Swissprot_40: VTA2_XENLA +	7.00	85.56	251.38	1807	I P18709 xenopus laevis (afri	Swissprot_40: YHGG_ECOLI +	6.00	88.76	3.9e+03	78	I P46845 escherichia coli, a
Swissprot_40: PKR2_CDICI +	7.00	85.37	250.62	1858	I P54674 dictyostellium discoi	Swissprot_40: C551_ECTHL +	6.00	88.67	3.9e+03	79	I P38587 ectothiorhodospira
Swissprot_40: TAGB_CDICI +	7.00	85.20	249.95	1905	I P54683 dictyostellium discoi	Swissprot_40: R117_APIME +	6.00	88.57	3.9e+03	79	I P15857 apis mellifera (hon
Swissprot_40: RPOB_PVMR +	7.00	84.97	249.09	1967	I P17965 potato virus m (strai	Swissprot_40: R117_APIME +	6.00	88.57	3.9e+03	80	I Q97qc0 streptococcus pneum
Swissprot_40: CYAA_YEAST +	7.00	84.77	248.29	2026	I P08678 saccharomyces cerevi	Swissprot_40: YNYC_STYCL -	6.00	88.58	3.9e+03	80	I Q18494 styela clava (sea s
Swissprot_40: FAS1_YEAST -	7.00	84.68	247.97	2051	I P07149 s fatty acid synthas	Swissprot_40: Y509_ECO57 -	6.00	88.58	3.9e+03	80	I P58092 escherichia coli (s
Swissprot_40: N214_HUMAN -	7.00	84.55	247.46	2090	I P35658 homo sapiens (human)	Swissprot_40: YHYS_WOLSU -	6.00	88.58	3.9e+03	80	I P31877 wolinnella succinoge
Swissprot_40: YRFL_SPIOI +	7.00	84.42	246.95	2131	I P08973 spinacia oleracea (s	Swissprot_40: YXIT_BACSU -	6.00	88.58	3.9e+03	80	I P42311 bacillus subtilis .
Swissprot_40: RRP2_SEOUB -	7.00	84.35	246.70	2151	I P27314 seoul virus (strain	Swissprot_40: R31B_LACIA -	6.00	88.49	3.9e+03	81	I Q9cf85 lactococcus lactis .
Swissprot_40: RRPPL_PUUMH +	7.00	84.33	246.64	2156	I P27176 puumala virus (strai	Swissprot_40: RPOH_THECE -	6.00	88.41	3.8e+03	82	I P31815 thermococcus celer .
Swissprot_40: POIG_PUUMH +	7.00	84.26	246.35	2179	I P03303 human rhinovirus 14	Swissprot_40: YBDJ_KLEPN +	6.00	88.41	3.8e+03	82	I Q48414 klebsiella pneumoni
Swissprot_40: RRPPL_NDVB +	7.00	84.18	246.05	2204	I P11205 newcastle disease vir	Swissprot_40: PGCA_PIG -	6.00	88.32	3.8e+03	83	I Q29011 sus scrofa (pig). agg
Swissprot_40: RRPPO_TACV +	7.00	84.16	245.98	2210	I P20430 tacaribe virus. rna p	Swissprot_40: GVM1_ECALNI -	6.00	88.24	3.8e+03	84	I P26377 halobacterium sp. (
Swissprot_40: THYG_HUMAN -	7.00	82.60	240.08	2768	I P01266 homo sapiens (human)	Swissprot_40: TAG1_ECOLI -	6.00	88.24	3.8e+03	84	I P74301 escherichia coli, t
Swissprot_40: PGBM_MOUSE +	7.00	80.56	231.34	3707	I Q05793 mus musculus (mouse)	Swissprot_40: Y073_NPVOP -	6.00	88.24	3.8e+03	84	I Q10326 oryia pseudotsugat
Swissprot_40: LMA5_MOUSE -	7.00	80.54	231.27	3718	I Q61001 mus musculus (mouse)	Swissprot_40: YSMA_SERVA +	6.00	88.16	3.8e+03	85	I P14307 serratia marcescens
Swissprot_40: SACS_HUMAN -	7.00	80.34	230.50	3829	I Q9nz14 homo sapiens (human)	Swissprot_40: ITP2_CHICK -	6.00	88.16	3.8e+03	85	I Q90683 gallus gallus (chic
Swissprot_40: SACS_MOUSE -	7.00	80.34	230.49	3830	I Q9j1c8 mus musculus (mouse)	Swissprot_40: GUPM_HALME -	6.00	88.08	3.8e+03	86	I Q02238 halobacterium medit
Swissprot_40: SID2_USTMA -	7.00	80.13	229.71	3947	I Q43103 ustilago maydis (smu	Swissprot_40: NVLM_ARTSA -	6.00	88.08	3.8e+03	86	I P19049 artemia salina (bri

SwissProt_40:R31B_STRPY -	6.00	88.08	3.8e+03	86	! Q9a016 streptococcus pyogenes	SwissProt_40:VNST_BUNGE +	6.00	86.43	3.7e+03	109	! P16992 bunyavirus germist
SwissProt_40:U186_HUMAN -	6.00	88.08	3.8e+03	86	! Q9bwj5 homo sapiens (human).	SwissProt_40:YF79_MYCPN +	6.00	86.43	3.7e+03	109	! P75201 mycoplasma pneumon
SwissProt_40:U186_MOUSE -	6.00	88.08	3.8e+03	86	! Q923d4 mus musculus (mouse).	SwissProt_40:YJ49_YEAST +	6.00	86.43	3.7e+03	109	! P47126 saccharomyces cere
SwissProt_40:RR15_TOBAC -	6.00	88.00	3.8e+03	87	! P56805 nicotiana tabacum (com	SwissProt_40:ZM31_RATZ +	6.00	86.43	3.7e+03	109	! O82106 zea mays (maize).
SwissProt_40:RR15_ARATH -	6.00	87.92	3.8e+03	88	! P56803 arabidopsis thaliana	SwissProt_40:AGPL_RATZ +	6.00	86.36	3.7e+03	110	! O35460 rattus norvegicus (r
SwissProt_40:YC29_BACHD -	6.00	87.92	3.8e+03	88	! Q8kd1a bacillus halodurans. h	SwissProt_40:RPCH_CARMA +	6.00	86.36	3.7e+03	110	! Q26324 carcinus maenas (c
SwissProt_40:DRPH_ROMM1 -	6.00	87.84	3.8e+03	89	! P09929 romalea microptera (lu	SwissProt_40:RSII_STRCI +	6.00	86.36	3.7e+03	110	! P29606 streptomyces cacao
SwissProt_40:DRPH_ROMM2 -	6.00	87.84	3.8e+03	89	! P34053 methanococcus jannasch	SwissProt_40:VFUS_VACCC +	6.00	86.36	3.7e+03	110	! P29535 vaccinia virus (st
SwissProt_40:RLJ4_METJA -	6.00	87.84	3.8e+03	89	! P08791 pseudomonas putida. 30	SwissProt_40:VFUS_VACCV +	6.00	86.36	3.7e+03	110	! P11258 vaccinia virus (st
SwissProt_40:RS15_PSEPU -	6.00	87.84	3.8e+03	89	! P09386 bacteriophages 933w. sp	SwissProt_40:VFUS_VAVV +	6.00	86.36	3.7e+03	110	! P33816 variola virus. 14 k
SwissProt_40:SLTB_BP933 -	6.00	87.84	3.8e+03	89	! P76195 escherichia coli. and	SwissProt_40:NG22_DROME +	6.00	86.24	3.7e+03	112	! P40139 drosophila melanoga
SwissProt_40:YDIH_ECOLI -	6.00	87.76	3.8e+03	89	! P33083 glycine max (soybean).	SwissProt_40:FRT2_HUMAN +	6.00	86.18	3.7e+03	113	! O75474 homo sapiens (huma
SwissProt_40:AX6B_SOYBN -	6.00	87.76	3.8e+03	90	! Q9mk9 mesostigma viride. chl	SwissProt_40:U139_CABEL +	6.00	86.18	3.7e+03	113	! Q09993 caenorhabditis ele
SwissProt_40:RR15_MESVI -	6.00	87.76	3.8e+03	90	! Q84864 lactobacillus sakei. s	SwissProt_40:Y109_HAETN +	6.00	86.18	3.7e+03	113	! P43943 haemophilus influ
SwissProt_40:SA1A_LACSK +	6.00	87.76	3.8e+03	90	! Q92d4a rickettsia prowazeki.	SwissProt_40:Y101_PARTE +	6.00	86.18	3.7e+03	113	! P15603 paramyxium tetrau
SwissProt_40:Y169_RICPR -	6.00	87.76	3.8e+03	90	! P55490 rhizobium sp. (strain	SwissProt_40:NU3M_ALLMA -	6.00	86.12	3.7e+03	114	! P37399 allomyces macrogyn
SwissProt_40:Y4IG_RHTSN +	6.00	87.76	3.8e+03	90	! O83921 treponema pallidum. by	SwissProt_40:RL12_HALHA -	6.00	86.12	3.7e+03	114	! P05768 halobacterium halo
SwissProt_40:Y955_TREPA +	6.00	87.76	3.8e+03	90	! Q04758 mus musculus (mouse).	SwissProt_40:ACPL_MYCTU -	6.00	86.06	3.7e+03	115	! Q10500 mycobacterium tuber
SwissProt_40:IPKB_MOUSE -	6.00	87.61	3.8e+03	92	! P46514 helianthus annuus (com	SwissProt_40:INS_MYXGL -	6.00	86.06	3.7e+03	115	! P01342 myxine glutinosa (a
SwissProt_40:LE10_HELAN -	6.00	87.61	3.8e+03	92	! P30523 schizosaccharomyces po	SwissProt_40:NU3M_BOVIN +	6.00	86.06	3.7e+03	115	! P03898 bos taurus (bovine
SwissProt_40:UCRQ_SCHPO +	6.00	87.61	3.8e+03	92	! P44282 haemophilus influenzae	SwissProt_40:RNPA_BUCAI -	6.00	86.06	3.7e+03	115	! P57130 buchnera aphidicol
SwissProt_40:YG51_HAETN +	6.00	87.61	3.8e+03	94	! P71996 mycobacterium tubercul	SwissProt_40:YDHD_ECOLI -	6.00	86.06	3.7e+03	115	! P37010 escherichia coli.
SwissProt_40:YH38_MYCTU -	6.00	87.46	3.8e+03	94	! Q05035 caenorhabditis elegans	SwissProt_40:PHP3_SCHPO +	6.00	85.99	3.7e+03	116	! P36611 schizosaccharomyce
SwissProt_40:YLA3_CABEL +	6.00	87.46	3.8e+03	94	! P11830 saccharopolyspora eryth	SwissProt_40:STP2_RAT +	6.00	85.99	3.7e+03	116	! P11011 rattus norvegicus (r
SwissProt_40:ACP_SACER -	6.00	87.38	3.8e+03	95	! P36690 streptomyces griseus.	SwissProt_40:GF62_HUMAN +	6.00	85.94	3.7e+03	117	! O08765 homo sapiens (huma
SwissProt_40:SCCE_STRGR -	6.00	87.38	3.8e+03	95	! P26879 legionella pneumophila	SwissProt_40:MTPN_HUMAN +	6.00	85.94	3.7e+03	117	! P58546 homo sapiens (huma
SwissProt_40:R101_LEGPN -	6.00	87.31	3.8e+03	96	! P47926 lycopersicon esculentu	SwissProt_40:MTPN_MOUSE +	6.00	85.94	3.7e+03	117	! P80144 mus musculus (mous
SwissProt_40:RS1L_LYCES -	6.00	87.31	3.8e+03	96	! P27004 xenopus laevis (africa	SwissProt_40:PRD1_CABEL -	6.00	85.94	3.7e+03	117	! Q17827 caenorhabditis ele
SwissProt_40:S110_XENLA -	6.00	87.31	3.8e+03	96	! Q9wval01 rattus norvegicus (rat)	SwissProt_40:RL22_SCHPO -	6.00	85.94	3.7e+03	117	! Q09668 schizosaccharomyce
SwissProt_40:IM8A_RAT +	6.00	87.24	3.8e+03	97	! P39565 saccharomyces cerevisi	SwissProt_40:MTPN_CHICK +	6.00	85.88	3.7e+03	118	! Q91955 gallus gallus (chl
SwissProt_40:YAN9_YEAST +	6.00	87.24	3.8e+03	97	! P46688 arabidopsis thaliana	SwissProt_40:YFPB_ECOLI +	6.00	85.88	3.7e+03	118	! P75111 mycoplasma pneumon
SwissProt_40:GAS1_ARATH +	6.00	87.17	3.8e+03	98	! P05507 rattus norvegicus (rat).	SwissProt_40:YF5C_ECOLI -	6.00	85.88	3.7e+03	118	! P19170 escherichia coli.
SwissProt_40:CNLM_STRCP -	6.00	87.17	3.8e+03	98	! Q03152 struthio camelus (ostr	SwissProt_40:ACLY_ACHLY -	6.00	85.82	3.7e+03	119	! P81730 achromobacter lyti
SwissProt_40:CNLM_STYNA +	6.00	87.10	3.8e+03	99	! P46688 arabidopsis thaliana	SwissProt_40:IP1_BACSU -	6.00	85.82	3.7e+03	119	! P33804 bacillus subtilis.
SwissProt_40:GAS2_ARATH +	6.00	87.10	3.8e+03	99	! P05114 homo sapiens (human).	SwissProt_40:PI15_RAT +	6.00	85.82	3.7e+03	119	! Q63396 rattus norvegicus (ra
SwissProt_40:HG14_HUMAN -	6.00	87.10	3.8e+03	99	! P31324 mus musculus (mouse).	SwissProt_40:PI15_ECOLI -	6.00	85.82	3.7e+03	119	! P14494 escherichia coli.
SwissProt_40:KAP3_MOUSE -	6.00	87.10	3.8e+03	99	! P28291 bos taurus (bovine).	SwissProt_40:WN1A_PLEJO -	6.00	85.82	3.7e+03	119	! P28131 plethodon jordani
SwissProt_40:MCPA_BOVIN -	6.00	87.10	3.8e+03	99	! P40932 myeloblastic leukaemic	SwissProt_40:CHH4_PENMO -	6.00	85.76	3.7e+03	120	! O97386 penaeus monodon (p
SwissProt_40:ENV6_MPLV -	6.00	87.03	3.8e+03	100	! P26931 lactobacillus ferment	SwissProt_40:RL18_TREPA -	6.00	85.76	3.7e+03	120	! O83235 treponema pallidum
SwissProt_40:URE3_LACFE -	6.00	87.03	3.8e+03	100	! Q9t157 bacteriophage apse-1.	SwissProt_40:RUS_ECOLI +	6.00	85.76	3.7e+03	120	! O40116 escherichia coli. c
SwissProt_40:VATN_HUMAN -	6.00	87.03	3.8e+03	100	! Q75787 homo sapiens (human).	SwissProt_40:YGLV_YEAST +	6.00	85.76	3.7e+03	120	! P53229 saccharomyces cere
SwissProt_40:Y151_ARCFU -	6.00	87.03	3.8e+03	100	! Q30086 archaeoglobus fulgidu	SwissProt_40:H2B1_PATGR +	6.00	85.70	3.7e+03	121	! P02284 patella granatina (
SwissProt_40:YUPO_YEAST -	6.00	87.03	3.8e+03	100	! P47004 saccharomyces cerevisi	SwissProt_40:HV01_MOUSE +	6.00	85.70	3.7e+03	121	! P01745 mus musculus (mous
SwissProt_40:YD2D_SCHPO -	6.00	86.96	3.8e+03	101	! Q10261 schizosaccharomyces p	SwissProt_40:RK14_CYACA -	6.00	85.70	3.7e+03	121	! Q9t1u2 cyanidium caldariu
SwissProt_40:YMDH_PENVA +	6.00	86.89	3.8e+03	102	! P53322 penaeus vannamei (penc	SwissProt_40:RLU7A_THCAC +	6.00	85.70	3.7e+03	121	! Q9hj56 thermoplasma acido
SwissProt_40:RS6_DEIRA -	6.00	86.89	3.8e+03	102	! Q9t52 deinococcus radioduran	SwissProt_40:FLIT_SALTY -	6.00	85.64	3.7e+03	122	! P26611 salmonella typhimu
SwissProt_40:VP21_BPAPS -	6.00	86.89	3.8e+03	102	! Q9t157 bacteriophage apse-1.	SwissProt_40:FOLB_CHLMI +	6.00	85.64	3.7e+03	122	! Q9pj66 chlamydia muridar
SwissProt_40:Y389_METJA +	6.00	86.89	3.8e+03	102	! Q58007 methanococcus jannasch	SwissProt_40:H2B1_PSAMEL -	6.00	85.64	3.7e+03	122	! P02287 psammecolus milla
SwissProt_40:Y494_CYAPA +	6.00	86.89	3.8e+03	102	! P15811 cyanophora paradoxa.	SwissProt_40:SAAL_MOUSE -	6.00	85.64	3.7e+03	122	! P05366 mus musculus (mous
SwissProt_40:YCU0_YEAST +	6.00	86.89	3.8e+03	102	! P25630 saccharomyces cerevisi	SwissProt_40:SAAT_MOUSE -	6.00	85.64	3.7e+03	122	! P05367 mus musculus (mous
SwissProt_40:RLA4_PLAFA +	6.00	86.82	3.8e+03	103	! Q97231 plasmodium falciparum	SwissProt_40:Y22B_MERTJ -	6.00	85.64	3.7e+03	122	! P81305 methanococcus jann
SwissProt_40:RR10_ODOSI +	6.00	86.82	3.8e+03	103	! P75074 mycoplasma pneumoniae	SwissProt_40:RS13_MYCGE +	6.00	85.59	3.7e+03	123	! P47421 mycoplasma genital
SwissProt_40:Y4A0_MYCPN +	6.00	86.82	3.8e+03	103	! Q9zm09 helicobacter pylori. 3	SwissProt_40:Y949_METJA +	6.00	85.59	3.7e+03	123	! Q38359 methanococcus jann
SwissProt_40:RL21_HELPJ +	6.00	86.75	3.7e+03	104	! P56046 helicobacter pylori. 4	SwissProt_40:APCA_RABIT +	6.00	85.53	3.7e+03	124	! P50557 oryctolagus cunicu
SwissProt_40:RMT2_BRAJA +	6.00	86.75	3.7e+03	104	! Q9rh70 bradyrhizobium japoni	SwissProt_40:FOLB_CHLTR -	6.00	85.53	3.7e+03	124	! O84620 chlamydia trachoma
SwissProt_40:SP03_ARATH -	6.00	86.75	3.7e+03	104	! P55852 arabidopsis thaliana	SwissProt_40:H2B1_CHITH +	6.00	85.53	3.7e+03	124	! P21897 chironomus thummi t
SwissProt_40:VATE_THCAC +	6.00	86.75	3.7e+03	104	! Q9hm66 thermoplasma acidophi	SwissProt_40:HM7X_CHICK +	6.00	85.53	3.7e+03	124	! P15142 gallus gallus (chl
SwissProt_40:Y4V4_XANCV -	6.00	86.75	3.7e+03	104	! P19519 xanthomonas campestr	SwissProt_40:YBF2_DROME +	6.00	85.53	3.7e+03	124	! Q9vnl3 drosophila melanog
SwissProt_40:Y032_HALN1 -	6.00	86.75	3.7e+03	104	! Q9hm43 halobacterium sp. (st	SwissProt_40:YAF3_YEAST +	6.00	85.53	3.7e+03	124	! P38190 saccharomyces cere
SwissProt_40:PDXA_ERWHE +	6.00	86.69	3.7e+03	105	! P47834 erwinia herbicola. 4-	SwissProt_40:ALR3_HUMAN +	6.00	85.47	3.7e+03	125	! P55789 homo sapiens (huma
SwissProt_40:GAS4_ARATH -	6.00	86.62	3.7e+03	106	! Q46690 arabidopsis thaliana	SwissProt_40:CL163_HUMAN +	6.00	85.47	3.7e+03	125	! Q95177 homo sapiens (huma
SwissProt_40:RL11_HAEDU -	6.00	86.62	3.7e+03	106	! Q32613 haemophilus ducreyi.	SwissProt_40:RL7_RICPR -	6.00	85.47	3.7e+03	125	! Q9ze21 rickettsia prowazek
SwissProt_40:RL11_HAEDU -	6.00	86.62	3.7e+03	106	! P05318 saccharomyces cerevisi	SwissProt_40:DOCR_HUMAN +	6.00	85.42	3.7e+03	126	! O75956 homo sapiens (huma
SwissProt_40:HSP2_MOUSE +	6.00	86.56	3.7e+03	107	! P07978 mus musculus (mouse).	SwissProt_40:PI15_HUMAN +	6.00	85.42	3.7e+03	126	! P53999 homo sapiens (huma
SwissProt_40:NG1_DROME +	6.00	86.56	3.7e+03	107	! P23938 drosophila melanogaste	SwissProt_40:PI15_MOUSE +	6.00	85.42	3.7e+03	126	! P11031 mus musculus (mouse
SwissProt_40:INS_BRARE -	6.00	86.49	3.7e+03	108	! Q73727 brachydanio rerio (zeb	SwissProt_40:PAHO_CAVPO -	6.00	85.42	3.7e+03	126	! P11083 cavia porcellus (g
SwissProt_40:RSGB_BACLI +	6.00	86.49	3.7e+03	108	! O50230 bacillus licheniformi	SwissProt_40:WNT1_ALOVU +	6.00	85.42	3.7e+03	126	! P28100 alopias vulpinus (
SwissProt_40:Y203_METJA -	6.00	86.49	3.7e+03	108	! Q60268 methanococcus jannasch	SwissProt_40:Y4SK_RHTSN +	6.00	85.42	3.7e+03	126	! P55654 rhizobium sp. (str
SwissProt_40:CYC6_CYACA +	6.00	86.43	3.7e+03	109	! Q9tlv1 cyanidium caldarium.	SwissProt_40:ILBP_RABIT -	6.00	85.36	3.7e+03	127	! P50119 oryctolagus cunicu
SwissProt_40:RL11_TRVCR -	6.00	86.43	3.7e+03	109	! P26643 trypanosoma cruzi. 1	SwissProt_40:K84T_UMV4 +	6.00	85.36	3.7e+03	127	! Q90121 ustilago maydis p4
SwissProt_40:RS17_HALHA -	6.00	86.43	3.7e+03	109	! Q24786 halobacterium halobid	SwissProt_40:LY6D_MOUSE +	6.00	85.36	3.7e+03	127	! P34559 mus musculus (mous

Swissprot_40:NIFW_RHILO +	6.00	85.36	3.7e+03	127	Q98at3 rhizobium loti (mesor	Swissprot_40:YDE6_SCHPO +	6.00	84.74	3.6e+03	139	Q10440 schizosaccharomyce
Swissprot_40:NKL_METTH -	6.00	85.36	3.7e+03	127	O26703 methanobacterium ther	Swissprot_40:PSBR_SPIOU -	6.00	84.69	3.6e+03	140	P10690 spinacia oleracea
Swissprot_40:SPK_PIG -	6.00	85.36	3.7e+03	127	P01359 sus scrofa (pig). spasm	Swissprot_40:RL1L_BACSU -	6.00	84.69	3.6e+03	140	P06796 bacillus subtilis
Swissprot_40:THIO_NEUCR -	6.00	85.36	3.7e+03	127	P42115 neurospora crassa. th	Swissprot_40:Y128_LISMO +	6.00	84.69	3.6e+03	140	P38702 listeria monocycog
Swissprot_40:CD59_AOTPR +	6.00	85.31	3.7e+03	128	P51447 aotus trivirgatus (nl	Swissprot_40:Y175_LISIN +	6.00	84.69	3.6e+03	140	Q92fd1 listeria innocua
Swissprot_40:CD59_CALSQ +	6.00	85.31	3.7e+03	128	P46657 callithrix sp. (marmos	Swissprot_40:YF85_XYLFA +	6.00	84.69	3.6e+03	140	Q99t1 xylella fastidiosa
Swissprot_40:CD59_BACSU -	6.00	85.31	3.7e+03	128	O08309 bacillus subtilis. sp	Swissprot_40:YF85_XYLFA +	6.00	84.69	3.6e+03	140	Q99t1 xylella fastidiosa
Swissprot_40:COTV_BACSU -	6.00	85.31	3.7e+03	128	O08309 bacillus subtilis. sp	Swissprot_40:RL1L_CHLMU -	6.00	84.64	3.6e+03	141	O9pk76 chlamydia muridar
Swissprot_40:IF5A_ARCFU +	6.00	85.31	3.7e+03	128	O29612 archaeoglobus fulgidu	Swissprot_40:RL1L_CHLMU -	6.00	84.64	3.6e+03	141	O9294 chlamydia pneumonia
Swissprot_40:RNP_PREEN +	6.00	85.31	3.7e+03	128	P19644 presbytis entellus (ha	Swissprot_40:RL1L_CHLTR -	6.00	84.64	3.6e+03	141	O84321 chlamydia trachom
Swissprot_40:SH2A_HUMAN +	6.00	85.31	3.7e+03	128	P06880 homo sapiens (human).	Swissprot_40:RL1L_ECOLI -	6.00	84.64	3.6e+03	141	P02409 escherichia coli,
Swissprot_40:YXZ2_YEAST -	6.00	85.31	3.7e+03	128	P25649 saccharomyces cerevis	Swissprot_40:RL1L_HAELN -	6.00	84.64	3.6e+03	141	P44331 haemophilus influe
Swissprot_40:ACP2_HORVU +	6.00	85.26	3.7e+03	129	P08817 hordeum vulgare (barl	Swissprot_40:RL1L_PROVU -	6.00	84.64	3.6e+03	141	P10055 proteus vulgaris.
Swissprot_40:FLCB_BORVU +	6.00	85.26	3.7e+03	129	P24500 bacillus subtilis. fl	Swissprot_40:RL1L_SERNA -	6.00	84.64	3.6e+03	141	P09763 serratia marcescen
Swissprot_40:IR08_HCMVA +	6.00	85.26	3.7e+03	129	P16806 human cytomegalovirus	Swissprot_40:YH11_SCHPO -	6.00	84.64	3.6e+03	141	O14177 schizosaccharomyce
Swissprot_40:RS1A_STRCU +	6.00	85.26	3.7e+03	129	P33095 strongylocentrotus pu	Swissprot_40:YH11_SCHPO -	6.00	84.59	3.6e+03	142	P14944 carcinus maenas (co
Swissprot_40:Y202_MYCGA +	6.00	85.26	3.7e+03	129	P06267 methanococcus jannasc	Swissprot_40:YH13_PYRAB +	6.00	84.59	3.6e+03	142	P00218 thermoplasma acidop
Swissprot_40:NIPE_MYCGA +	6.00	85.20	3.7e+03	130	P33255 mycoplasma gallisepti	Swissprot_40:FER_THEAC +	6.00	84.59	3.6e+03	142	P48221 coregonus autumnal
Swissprot_40:GUN2_PERAIE -	6.00	85.20	3.7e+03	130	P23666 persea americana (avo	Swissprot_40:GTH2_CORAU +	6.00	84.59	3.6e+03	142	Q92195 agarticus bisporus
Swissprot_40:SPFH_THEMA -	6.00	85.20	3.7e+03	130	Q92C3 thermotoga maritima.	Swissprot_40:GTH2_CORAU +	6.00	84.59	3.6e+03	142	Q92195 agarticus bisporus
Swissprot_40:TAT_HV2BE -	6.00	85.20	3.7e+03	130	P18098 human immunodeficien	Swissprot_40:RL1L_MYCTU -	6.00	84.59	3.6e+03	142	O9cbk1 mycobacterium lepr
Swissprot_40:TAT_HV2G1 -	6.00	85.20	3.7e+03	130	P17759 human immunodeficien	Swissprot_40:YB71_YEAST +	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2G1 -	6.00	85.20	3.7e+03	130	P18044 human immunodeficien	Swissprot_40:YB71_YEAST +	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2KR -	6.00	85.20	3.7e+03	130	Q74124 human immunodeficien	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2N2 -	6.00	85.20	3.7e+03	130	P05909 human immunodeficien	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2RO -	6.00	85.20	3.7e+03	130	P04605 human immunodeficien	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2SB -	6.00	85.20	3.7e+03	130	P12453 human immunodeficien	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2ST -	6.00	85.20	3.7e+03	130	P20880 human immunodeficien	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:TAT_HV2ST -	6.00	85.20	3.7e+03	130	P20880 human immunodeficien	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:YNO3_YEAST -	6.00	85.20	3.7e+03	130	P53908 saccharomyces cerevis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:ACSW_VULVU -	6.00	85.15	3.7e+03	131	P17940 vulpes vulpes (red fo	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:ACSW_VULVU -	6.00	85.15	3.7e+03	131	P17940 vulpes vulpes (red fo	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.6e+03	143	P35383 rhizobium sp. (str
Swissprot_40:RSGF_KLEPN +	6.00	85.15	3.7e+03	131	P09305 coregonus autumnalis	Swissprot_40:YK11_MYCTU -	6.00	84.54	3.		



SwissProt_40:YB29_MYCPN -	6.00	84.25	3.6e+03	149	! P75346 mycoplasma pneumoniae	SwissProt_40:HPPK_HAEIN -	6.00	83.76	3.6e+03	160	! P43777 haemophilus influ
SwissProt_40:YB56_LISMO -	6.00	84.25	3.6e+03	149	! P75346 mycoplasma pneumoniae	SwissProt_40:HTF4_PAPHA -	6.00	83.76	3.6e+03	160	! Q28772 papio hamadryas
SwissProt_40:Y547_CABEL -	6.00	84.25	3.6e+03	149	! Q09373 caenorhabditis elegans	SwissProt_40:ISPE_HAEDU +	6.00	83.76	3.6e+03	160	! Q47956 haemophilus ducrey
SwissProt_40:ABP5_MATZE -	6.00	84.21	3.6e+03	150	! P33489 zea mays (maize). aux	SwissProt_40:MOAC_HAEIN -	6.00	83.76	3.6e+03	160	! P45310 haemophilus influ
SwissProt_40:APDV_SALTY +	6.00	84.21	3.6e+03	150	! Q9xam6 salmonella typhimurium	SwissProt_40:PETD_SVNP2 -	6.00	83.76	3.6e+03	160	! P28057 synecococcus sp.
SwissProt_40:RNS_BOVIN +	6.00	84.21	3.6e+03	150	! P00669 bos taurus (bovine). h	SwissProt_40:PHAA_AGLNE +	6.00	83.76	3.6e+03	160	! P28555 aglaothamnion negl
SwissProt_40:VSL3_ARATH +	6.00	84.21	3.6e+03	150	! P49203 arabidopsis thaliana	SwissProt_40:PHAA_ANACP +	6.00	83.76	3.6e+03	160	! P07325 anabaena cylindric
SwissProt_40:VPL_BHPH1 -	6.00	84.21	3.6e+03	150	! P51722 bacteriophage hpi. prc	SwissProt_40:PHAA_ANASP +	6.00	83.76	3.6e+03	160	! P08055 anabaena sp. (stra
SwissProt_40:YB59_BPP2 -	6.00	84.21	3.6e+03	150	! P36934 bacteriophage p2. tail	SwissProt_40:PHAA_CVAPA +	6.00	83.76	3.6e+03	160	! P00316 cyanophora paradox
SwissProt_40:YB9L_YEAST -	6.00	84.21	3.6e+03	150	! P38343 saccharomyces cerevis	SwissProt_40:PHAA_FREDI +	6.00	83.76	3.6e+03	160	! P15370 fremyella diplosip
SwissProt_40:YC41_AQUAE -	6.00	84.21	3.6e+03	150	! Q67286 aquifex aeolicus. hyf	SwissProt_40:PHAA_MASLA +	6.00	83.76	3.6e+03	160	! P00315 mastigella diplosip
SwissProt_40:YH53_SCHPO +	6.00	84.21	3.6e+03	150	! P060154 schizosaccharomyces b	SwissProt_40:PHAA_PORPU +	6.00	83.76	3.6e+03	160	! P00315 mastigella diplosip
SwissProt_40:YHNR_CAPCA +	6.00	84.16	3.6e+03	151	! P79351 capreolus capreolus b	SwissProt_40:PHAA_SPIPL +	6.00	83.76	3.6e+03	160	! P51262 porphyra purpurea
SwissProt_40:TCGR_VIBCH +	6.00	84.16	3.6e+03	151	! P29483 vibrio cholerae. tox	SwissProt_40:PHAA_SPIPL +	6.00	83.76	3.6e+03	160	! P72504 spirulina purpurea
SwissProt_40:YG52_ANACE +	6.00	84.16	3.6e+03	151	! P35664 anaplasma centrale. h	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! P30176 escherichia coli.
SwissProt_40:YD28_AQUAE -	6.00	84.16	3.6e+03	151	! Q66419 aquifex aeolicus. hyf	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:YD28_CAEEL +	6.00	84.11	3.6e+03	152	! Q66419 aquifex aeolicus. hyf	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:EXBB_PASHA +	6.00	84.11	3.6e+03	152	! Q77749 caenorhabditis elegans	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:EXBB_PASHA +	6.00	84.11	3.6e+03	152	! Q77749 caenorhabditis elegans	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:HTF4_XENLA -	6.00	84.11	3.6e+03	152	! P07202 pasteurella haemolytic	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00	84.11	3.6e+03	152	! Q91605 xenopus laevis (afriid	SwissProt_40:YB57_CAEEL -	6.00	83.76	3.6e+03	160	! Q09665 caenorhabditis ele
SwissProt_40:IL13_HUMAN +	6.00</										

```

SwissProt_40:YF70_HAEIN - 6.00 83.38 3.6e+03 169 | P44259 haemophilus influenza
SwissProt_40:C550_SYNPI + 6.00 83.34 3.6e+03 170 | Q5210 synechococcus sp. (st
SwissProt_40:CH18_DROV2 + 6.00 83.34 3.6e+03 170 | P24515 dirosophila virilis (d
SwissProt_40:COPT_ARATH + 6.00 83.34 3.6e+03 170 | Q39065 arabidopsis thaliana
SwissProt_40:VEAR_HCMA + 6.00 83.34 3.6e+03 170 | P09694 human cytomegalovirus
SwissProt_40:CFUA_HSV2 - 6.00 83.30 3.6e+03 171 | Q66674 equine herpesvirus ty
SwissProt_40:RS4_AERPE - 6.00 83.30 3.6e+03 171 | Q95B58 aeropyrum pernix. 30s
SwissProt_40:RT25_MOUSE + 6.00 83.30 3.6e+03 171 | Q9d125 mus musculus (mouse).
SwissProt_40:TOXS_VIBPA - 6.00 83.30 3.6e+03 171 | Q05939 vibrio parahaemolytic
SwissProt_40:DCD_CLOHI + 6.00 83.25 3.6e+03 172 | Q9znj8 clostridium histolytic
SwissProt_40:NRDR_STRCL + 6.00 83.25 3.6e+03 172 | Q86848 streptomyces clavulig
SwissProt_40:SSRD_MOUSE - 6.00 83.25 3.6e+03 172 | Q62186 mus musculus (mouse).
SwissProt_40:YF57_PYAB + 6.00 83.25 3.6e+03 172 | Q9uza8 pyrococcus abyssi. hy

seq_name: SwissProt_40:TLR7_HUMAN

seq_documentation_block:
ID TLR7_HUMAN STANDARD; PRT; 1049 AA.
AC Q9NYK1; Q9NR98;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE TLR7
DE Toll-like receptor 7 precursor.
GN TLR7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20477806; PubMed=11022119;
RA Du X., Poltorak A., Wei Y., Beutler B.;
RT "Three novel mammalian Toll-like receptors: gene structure,
RT expression, and evolution."
RL Eur. Cytokine Netw. 11:362-371(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20477807; PubMed=11022120;
RA Chuang T.-H., Ulevitch R.J.;
RT "Cloning and characterization of a sub-family of human Toll-like
RT receptors: hTLR7, hTLR8 and hTLR9."
RL Eur. Cytokine Netw. 11:372-378(2000).
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response (By
CC similarity).
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: Detected in brain, placenta, spleen, stomach,
CC small intestine, lung and in plasmacytoid pre-dendritic cells.
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -!- SIMILARITY: CONTAINS 27 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF240467; AAF60188.1; -.
CC EMBL; AF245702; AAF78035.1; -.
CC MIM; 300365; -.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR_Cterm.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003592; LRR_out.
CC InterPro; IPR003591; LRR_typ.

```

```

DR InterPro; IPR000157; TIR.
DR Pfam; PF00560; LRR; 12.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01582; TIR; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 4.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 3.
DR SMART; SM00255; TIR; 1.
DR PROSITE; PS50104; TIR; 1.
DR Receptor; Immune response; Inflammatory response; signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 26
FT CHAIN 27 1049
FT DOMAIN 27 839
FT TRANSMEM 840 860
FT DOMAIN 861 1049
FT REPEAT 43 64
FT REPEAT 65 87
FT REPEAT 110 126
FT REPEAT 127 149
FT REPEAT 151 170
FT REPEAT 171 195
FT REPEAT 203 226
FT REPEAT 228 247
FT REPEAT 248 275
FT REPEAT 289 312
FT REPEAT 314 337
FT REPEAT 339 368
FT REPEAT 369 392
FT REPEAT 396 419
FT REPEAT 421 443
FT REPEAT 492 515
FT REPEAT 516 540
FT REPEAT 541 564
FT REPEAT 566 588
FT REPEAT 595 618
FT REPEAT 619 644
FT REPEAT 649 672
FT REPEAT 674 697
FT REPEAT 698 721
FT REPEAT 723 745
FT REPEAT 746 769
FT REPEAT 772 795
FT DOMAIN 889 1036
FT CARBOHYD 66 66
FT CARBOHYD 69 69
FT CARBOHYD 167 167
FT CARBOHYD 202 202
FT CARBOHYD 215 215
FT CARBOHYD 361 361
FT CARBOHYD 413 413
FT CARBOHYD 488 488
FT CARBOHYD 523 523
FT CARBOHYD 534 534
FT CARBOHYD 590 590
FT CARBOHYD 679 679
FT CARBOHYD 720 720
FT CARBOHYD 799 799
FT CARBOHYD 725 725
FT CONFLICT 738 738
FT CONFLICT 738 738
SQ SEQUENCE 1049 AA; 120921 MW; 8C701E9E437F2721 CRC64;

```

alignment\_scores:

Quality: 1049.00  
Ratio: 1.000Length: 1049  
Gaps: 0  
Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x TLR7\_HUMAN

Align seg 1/1 to: TLR7\_HUMAN from: 1 to: 1049

```
85 ATGGTGTTTCCCAATGGGACACTGAAGACAGCAAAATCTTATCCTTTTAA 134
|||||
1 MetValPheProMetTrpThrLeuLysArgGlnIleLeuIleLeuPheAs 17
|||||
135 CATAAATCCTAAATTCCAAACCTCTTGGGGCTAGATGGTTTCTTAAACATC 184
|||||
17 ntlelleLeuIleSerLysLeuLeuGlyAlaArgTrpPheProLysThrL 34
|||||
185 TGCCCTGTGATGTCACCTCTGGAGTGTCCAAAGAACCATGTGATCGTGGAC 234
|||||
34 euProCysAspValThrLeuAspValProLysAsnHisValIleValasp 50
|||||
235 TGCACAGACAAGCATTTGCACAGAAATTCCTGGAGGTATTCACCAAGAACAC 284
|||||
51 CysThrAspLysHisLeuThrGluIleProGlyGlyIleProThrAsnTh 67
|||||
285 CAGGAACCTCACCTCACCATTAACACATACACAGACATCTCCCGAGGT 334
|||||
67 rThrAsnLeuThrLeuThrIleAsnHisIleProAspIleSerProAla 84
|||||
335 CCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGATGCAACTGT 384
|||||
84 erPheHisargLeuAspHisLeuValGluIleAspPheargCysasnCys 100
|||||
385 GTACCTATTCCACTGGGGTCAAAAACACATGTGCATCAAGAGGCTGCA 434
|||||
101 ValProIleProLeuGlySerLysAsnAsnMetCysIleLysArgGlu 117
|||||
435 GATTAAACCCAGAAGCTTTAGTGGACTCACTTATTTAAATCCCTTTACC 484
|||||
117 ntleLysProArgSerPheSerGlyLeuThrTyrLeuLysSerLeuTyrL 134
|||||
485 TGGATGGAAACAGCTACTAGAGATACCGCAGGCGCTCCCGCTAGCTTA 534
|||||
134 euAspGlyAsnGlnLeuLeuGluIleProGlnGlyLeuProProSerLeu 150
|||||
535 CAGCTTCTCAGCGCTTGAGGCCAACACATCTTTTCCATCAGAAAAGAGAA 584
|||||
151 GlnLeuLeuSerLeuGluAlaAsnAsnIlePheSerIleArgLysGluAs 167
|||||
585 TCTACAGAACCTGGCCACATAGAAATCTCTACTCTGGGCCAAACACTGTT 634
|||||
167 nLeuThrGluLeuAlaAsnIleGluIleLeuTyrLeuGlyGlnAsnCyst 184
|||||
635 ATTATCGAAATCCTTGTATGTTTCATATTCAATAGAGAAAGATGCCTTC 684
|||||
184 yTyrArgAsnProCysTyrValSerTyrSerIleGluLysAspAlaPhe 200
|||||
685 CTAAACTTGACAAAGTTAAAGTGTCTCTCCCTGAAAGATAACAATGTCAC 734
|||||
201 LeuAsnLeuThrLysLeuLysValLeuSerLeuLysAspAsnAsnValTh 217
|||||
735 AGCGTGCCCTACTGTTTGGCAATCTACTTTACAGAACTATATCTCTACA 784
|||||
217 rAlaValProThrValLeuProSerThrLeuThrGluLeuTyrLeuTyrA 234
|||||
785 ACAACATGATTGCAAAAATCCAAAGAGATGATTTTAATAACCTCAACCAA 834
|||||
234 snAsnMetIleAlaLysIleGlnGluAspAspPheAsnAsnLeuAsnGln 250
|||||
835 TTACAAATCTTTGACCTAAGTGAATTTGCCCTCGTGTGTATATAATGCCCC 884
|||||
251 LeuGlnIleLeuAspLeuSerGlyAsnCysProArgCysTyrAsnAlaPr 267
|||||
885 ATTTCCTTGTGGCGGTGTAATAAATAATTTCTCCCTACAGATCCCTGTAA 934
|||||
267 oPheProCysAlaProCysLysAsnAsnSerProLeuGlnIleProValA 284
|||||
935 ATGCTTTTTCATCGCTGACAGAAATTAAGTTTTTACGCTACACAGTAC 984
|||||
284 snAlaPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300
|||||
```

```
985 TCTCTTCAGCATGTGCCCCCAAGATGGTTTAAAGAACATCAACAAACTCCA 1034
|||||
301 SerLeuGlnHisValProProArgTrpPheLysAsnIleAsnLysLeuGl 317
|||||
1035 GGAATCGGATCTGTCCCAAACTTCTTGCCCAAGAAATTTGGGGATGCTA 1084
|||||
317 nGluLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaL 334
|||||
1085 AATTTCTGCATTTTCTCCACAGCCTCATCAATTTGGATCTGTCTTCAAT 1134
|||||
334 ysPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350
|||||
1135 TTTGAACTTCAGGTCTATCGTCATCTATGAATCTATCACAAAGCATTTTC 1184
|||||
351 PheGluLeuGlnValTyrArgAlaSerMetAsnLeuSerGlnAlaPheSe 367
|||||
1185 TTCACGTGAAAAGCCTGAAAATTTCTCGGGATCAGAGGATATGTCTTTAAG 1234
|||||
367 rSerLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysG 384
|||||
1235 AGTTCAAAAGCTTTAACTCTCGCCATTCACATAATCTTCAAAATCTTGAA 1284
|||||
384 LuLeuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400
|||||
1285 GTTCTTGATCTTTGGCACATAACTTTATAAAATTTGCTAACCTCAGCATGT 1334
|||||
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417
|||||
1335 TAAACAATTTAAAAGCTGAAAGTCATAGATCTTTCAGTGAATAAATAT 1384
|||||
417 euLysGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIleS 434
|||||
1385 CACCTTCAGGAGATTCAGTGAAGTTGGCTTCTGCTCAAAATGCCAGACT 1434
|||||
434 erProSerGlyAspSerSerGluValGlyPheCysSerAsnAlaArgThr 450
|||||
1435 TCTGTAGAAAGTTATGAACCCAGGTCCTCGGAACAAATTTACATTTATTCA 1484
|||||
451 SerValGluSerTyrGluProGlnValLeuGluGlnLeuHisTyrPhear 467
|||||
1485 ATATGATAGTATGCAAGAGTTGCAGATTCAAAACAAGAGGCTTCTT 1534
|||||
467 gTyrAspLysTyrAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484
|||||
1535 TCATGTCTGTTAATGAAAGCTGCTACAGTATGGCAGACCTTGGATCTA 1584
|||||
484 heMetSerValAsnGluSerCysTyrLysTyrGlyGlnThrLeuAspLeu 500
|||||
1585 AGTAAAAATAGTATATTTTGTCAAGTCCTCTGATTTTTCAGCATCTTTC 1634
|||||
501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517
|||||
1635 TTTCTCAAAATGCCATCTCTCAGGAAATCTCATTTAGCCAAACTCTTA 1684
|||||
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534
|||||
1685 ATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTTGGACTTCTCC 1734
|||||
534 snGlySerGluPheGlnProLeuAlaGluLeuArgTyrLeuAspPheSer 550
|||||
1735 AACAAACCGGCTTTGATTTACTCCTATTCAACAGCATTTTGAGAGCTTCA 1784
|||||
551 AsnAsnArgLeuAspLeuLeuHisSerThrAlaPheGluGluLeuHisLy 567
|||||
1785 ACTGGAAGTCTCGGATATAAGCAGTAATAGCCATTTATTTTCAATCAGAAG 1834
|||||
567 sleuGluValLeuAspIleSerSerAsnSerHisTyrPheGlnSerGluG 584
|||||
1835 GAATTACTCATATGCTAACTTTTACCAGAACTTAAAGGTTCTGCGAGAAA 1884
|||||
584 lyIleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600
|||||
```

1885 CTGATGATGAACACAAATGACATCTCTCTCCACAGCAGGACCATGGA 1934  
|||||  
601 LeuMetMetAsnAspAsnAspSerSerSerThrSerArgThrMetcl 617  
1935 GAGTGTCTCTTAGAACCTCTGGAATCTAGAGAAATCACTTAGATGTTT 1984  
|||||  
617 uSerGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspValL 634  
1985 TATGGAGAGAAGGTGATAACAGATCTTACAATATTCAAGAATCTGCTA 2034  
|||||  
634 euTrpArgGluGlyAspAsnArgTyLeuGlnLeuPheLysAsnLeuLeu 650  
2035 AAATTAGAGGAATTAGACATCTCTAAAAATCCCTAAAGTTCTTGCTTC 2084  
|||||  
651 LysLeuGluGluLeuAspLysSerLysAsnSerLeuSerPheLeuProse 667  
2085 TGGAGTTTGTGATGGTATCCCTCCAATCTAAAGAAATCTCTCTTTGGCCA 2134  
|||||  
667 rGlyValPheAspGlyMetProProAsnLeuLysAsnLeuSerLeuAlaL 684  
2135 AAATGGGCTCAATCTTTTCAGTTGGAAGAACTCCAGTGTCTAAAGAAC 2184  
|||||  
684 yAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
2185 CTGCAAACTTTGGACCTCAGCCACAACTGACCACTGTCCCTCAGAG 2234  
|||||  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrThrValProGluAr 717  
2235 ATTATCCAATCTGTTCAGAGGCTCAAGAATCTGATTTCTTAAGAATAATC 2284  
|||||  
717 gLeuSerAsnCysSerArgSerLeuLysAsnLeuLysAsnAsnG 734  
2285 AAATCAGGAGCTGACGAAGTATTTCTACAGATCCCTTCCAGTTGCCA 2334  
|||||  
734 InLeArgSerLeuThrLysTyPheLeuGlnAspAlaPheGlnLeuArg 750  
2335 TATCTGGATCTCAGCTCAATAAAATCCAGATGATCCAAAGACCAAGCTT 2384  
|||||  
751 TyrLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
2385 CCAGAAAATGTCCCTCAACAATCTGAAGATGTGTGCTTTTGCATCATATC 2434  
|||||  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuLeuHisAsnA 784  
2435 GGTTCCTGTCACCTGTGATGCTGTCTGTCTGTCTGTGGGTAAACCAT 2484  
|||||  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpValAsnHis 800  
2485 ACGGAGGTGACTATTCTTACCTTGGCCACAGATGTGACTTGTGTGGGGCC 2534  
|||||  
801 ThrGluValThrIleProTyLeuAlaThrAspValThrCysValGlyPr 817  
2535 AGGAGCACACAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
|||||  
817 oGlyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyThrCysG 834  
2585 AGTTAGATCTGACTAACCTGATCTCTCTCACTTCCATATCTGATCT 2634  
|||||  
834 lueLeuAspLeuThrAsnLeuIleLeuPheSerLeuSerLysValSer 850  
2635 CTCTTCTCATGTGTGATGATGACAGCAAGTCACTCTATTCTTGGGATGT 2684  
|||||  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyThrTrpaspVa 867  
2685 GTCGTATATTACCATTTCTGTAAAGCCCAAGATAAAGGGGTATCAGCGTC 2734  
|||||  
867 lTrpTyIleTyHisPheCysLysAlaLysIleLysGlyTyThrGlnArgL 884  
2735 TAATATCACAGACTGTGCTGATGATGCTTTATTTGTTGATGATGACATAA 2784  
|||||  
884 eulleSerProAspCysCysTyTrpAspAlaPheIleValTyAspThrLys 900  
2785 GACCCAGCTGTGACCGAGTGGGTTTTTGGCTGAGCTGGTGGCCAAACTGGA 2834

|||||  
901 AspProAlaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuG1 917  
2835 AGACCAAGAGAGAGAAACATTTTAAATTTATGCTCGAGGAAAGGGACTGGT 2884  
|||||  
917 uAspProArgGluLysHisPheAsnLeuCysLeuGluGluArgAspTrpL 934  
2885 TACCAGGGCAGCCAGTCTTGAAAAACCTTTCCAGAGATACAGCTTAGC 2934  
|||||  
934 euProGlyGlnProValLeuGluAsnLeuSerGlnSerIleGlnLeuSer 950  
2935 AAAAAGACAGTGTTTTGTGATGACAGACAAAGTATGCAAGACTGAAAAATTT 2984  
|||||  
951 LysLysThrValPheValMetThrAspLysTyAlaLysThrGluAsnPh 967  
2985 TAAGTAGACATTTTACTTGTCCCATCAGAGCTCATGGATGAAAAAGTTG 3034  
|||||  
967 eLysIleAlaPheTyLeuSerHisGlnArgLeuMetAspGluLysValA 984  
3035 ATCTGATTTATCTGATATTTCTTGAGAGCCCTTTCCAGAAAGTCCAAGTTC 3084  
|||||  
984 sPValIleLeuLeuPheLeuGluLysProPheGlnLysSerLysPhe 1000  
3085 CTCCAGCTCCGGAAAAAGGCTCTGTGGAGTCTGTCTTGTAGTGCCCAAC 3134  
|||||  
1001 LeuGlnLeuArgLysArgLeuCysGlySerSerValLeuGluTrpProth 1017  
3135 AAACCCGCAAGCTCACCATCTTCTTGGCAGTGTCTAAAGAACGCCCTGG 3184  
|||||  
1017 rAsnProGlnAlaHisProTyPheTrpGlnCysLeuLysAsnAlaLeuA 1034  
3185 CCACAGACAATCATGTGGCTATAGTCAGGTGTCTCAAGGAAACGGTC 3231  
|||||  
1034 lathrAspAsnHisValAlaLysSerGlnValPheLysGluThrVal 1049  
seq\_name: SwissProt\_40:TLR7\_MOUSE  
seq\_documentation\_block:  
ID TLR7\_MOUSE STANDARD; PRT; 1050 AA.  
AC P58681;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Toll-like receptor 7 precursor.  
GN TLR7.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Macrophage;  
RA Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;  
RT "Molecular cloning of murine Toll-Like-Receptor 7."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Participates in the innate immune response to microbial  
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B  
CC activation, cytokine secretion and the inflammatory response (By  
CC similarity).  
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -!- SIMILARITY: CONTAINS 28 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```

CC DR EMBL; AY035899; AAK62676.1; -
DR PROSITE; PS0104; TIR; 1.
KW Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1050 TOLL-LIKE RECEPTOR 7.
FT DOMAIN 27 837 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 838 858 POTENTIAL.
FT DOMAIN 859 1050 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 64 LRR 1.
FT REPEAT 65 87 LRR 2.
FT REPEAT 89 111 LRR 3.
FT REPEAT 126 149 LRR 4.
FT REPEAT 151 170 LRR 5.
FT REPEAT 171 195 LRR 6.
FT REPEAT 203 226 LRR 7.
FT REPEAT 228 247 LRR 8.
FT REPEAT 248 273 LRR 9.
FT REPEAT 275 289 LRR 10.
FT REPEAT 290 312 LRR 11.
FT REPEAT 314 337 LRR 12.
FT REPEAT 339 364 LRR 13.
FT REPEAT 369 392 LRR 14.
FT REPEAT 396 419 LRR 15.
FT REPEAT 421 443 LRR 16.
FT REPEAT 493 516 LRR 17.
FT REPEAT 517 542 LRR 18.
FT REPEAT 543 565 LRR 19.
FT REPEAT 567 589 LRR 20.
FT REPEAT 596 619 LRR 21.
FT REPEAT 620 645 LRR 22.
FT REPEAT 650 673 LRR 23.
FT REPEAT 675 698 LRR 24.
FT REPEAT 699 722 LRR 25.
FT REPEAT 724 746 LRR 26.
FT REPEAT 747 770 LRR 27.
FT REPEAT 773 796 LRR 28.
FT DOMAIN 890 1037 TIR.
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 524 524 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 535 535 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 680 680 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 800 800 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1050 AA; 121836 MW; 495B75DEE849D8EE CRC64;

```

```

alignment_scores:
  Quality: 48.00      Length: 48
  Ratio: 1.000        Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

```

alignment\_block:

US-09-202-054-2 x TLR7\_MOUSE ..

Align seg 1/1 to: TLR7\_MOUSE from: 1 to: 1050

```

2815 GAGCTGGTGGCCAACTGGAAGACCCAGAGAGAAACATTTTAATTATG 2864
|||||
912 GluLeuValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCy 928
|||||
2865 TCTCAGGAAAGGGACTGGTTACCAGGGCAGCGAGTCTCGGAAACCTTT 2914
|||||
928 sleuGluLysArgAspTrpLeuProGlyGlnProValLeuGluAsnLeus 945
|||||
2915 CCCAGAGCATACAGCTAGCAAAAAGACAGCTGTTTGTGATGACA 2958

```

```

|||||
945 erGlnSerIleGlnLeuSerLysThrValPheValMetThr 959
seq_name: SwissProt_40:TLR8_MOUSE
seq_documentation_block:
ID TLR8_MOUSE STANDARD; PRT; 1032 AA.
AC P58682;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Toll-like receptor 8 precursor.
GN TLR8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RA Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;
RT "Molecular cloning of murine Toll-like Receptor 8.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response (By
CC similarity).
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -!- SIMILARITY: CONTAINS 26 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AY035899; AAK62677.1; -
DR PROSITE; PS0104; TIR; 1.
DR Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 1032 TOLL-LIKE RECEPTOR 8.
FT DOMAIN 24 818 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 819 839 POTENTIAL.
FT DOMAIN 840 1032 CYTOPLASMIC (POTENTIAL).
FT REPEAT 41 61 LRR 1.
FT REPEAT 62 85 LRR 2.
FT REPEAT 87 109 LRR 3.
FT REPEAT 120 143 LRR 4.
FT REPEAT 145 165 LRR 5.
FT REPEAT 166 194 LRR 6.
FT REPEAT 195 218 LRR 7.
FT REPEAT 220 239 LRR 8.
FT REPEAT 240 267 LRR 9.
FT REPEAT 281 304 LRR 10.
FT REPEAT 306 329 LRR 11.
FT REPEAT 331 360 LRR 12.
FT REPEAT 361 384 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 436 LRR 15.
FT REPEAT 471 494 LRR 16.
FT REPEAT 520 543 LRR 17.
FT REPEAT 545 572 LRR 18.
FT REPEAT 574 598 LRR 19.
FT REPEAT 600 621 LRR 20.
FT REPEAT 629 652 LRR 21.
FT REPEAT 654 677 LRR 22.

```

```
FT REPEAT 678 701 LRR 23.
FT REPEAT 702 725 LRR 24.
FT REPEAT 727 749 LRR 25.
FT REPEAT 752 776 LRR 26.
FT DOMAIN 869 1016 TIR.
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 502 502 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 581 581 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 743 743 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1032 AA; 119338 MW; B2812191D99EE42D CRC64;

alignment_scores:
  Quality: 17.00 Length: 17
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x TLR8_MOUSE ..
Align seg 1/1 to: TLR8_MOUSE from: 1 to: 1032
844 CTTGACCTAAGTGGAAATGCCCTCGTGTATATGATGCCCATTTCCCTTG 893
246 LeuAspLeuSerGlyAsnCysProArgCysTyrAsnAlaProPheProCys 262
894 T 894
262 s 262

seq_name: SwissProt_40:TLR9_HUMAN
seq_documentation_block:
ID TLR9_HUMAN STANDARD; PRT: 1032 AA.
AC Q9NR96; Q9NYC3; Q9NYC2; Q9HD70; Q9HD69; Q9HD68;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Toll-like receptor 9 precursor.
GN TLR9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE-Monocytic leukemia;
RX Du X., Potorak A., Wei Y., Beutler B.;
RT "Three novel mammalian Toll-like receptors: gene structure,
RT expression, and evolution."
RL Eur. Cytokine Netw. 11:362-371(2000).
RN [2].
RP SEQUENCE FROM N.A. (ISOFORMS 1; 3; 4 AND 5).
RC TISSUE-Placenta;
RX PubMed-11022120;
```

```
RA Chuang T.-H., Ulevitch R.J.;
RT "Cloning and characterization of a sub-family of human Toll-like
RT receptors: hTLR7, hTLR8 and hTLR9.";
RL Eur. Cytokine Netw. 11:372-378(2000).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=1012018; PubMed=11130078;
RA Hemmi H., Takeuchi O., Kawai T., Kaisho T., Sato S., Sanjo H.,
RA Matsumoto M., Hoshino K., Wagner H., Takeda K., Akira S.;
RT "A Toll-like receptor recognizes bacterial DNA.";
RL Nature 408:740-745(2000).
[4]
RN FUNCTION.
RX PubMed=11564765;
RA Takeshita F., Leifer C.A., Gursel I., Ishii K.J., Takeshita S.,
RA Gursel M., Kilman D.M.;
RT "Cutting edge: role of Toll-like receptor 9 in CpG DNA-induced
RT activation of human cells.";
RL J. Immunol. 167:3555-3558(2001).
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. Detects the unmethylated cytidine-phosphate-guanosine
CC (cpG) motifs present in bacterial DNA. Acts via MyD88 and TRAF6,
CC leading to NF-kappa-B activation, cytokine secretion and the
CC inflammatory response.
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: 1/A (SHOWN HERE), 2/B,
CC 3, 4 AND 5; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: Highly expressed in spleen, lymph node, tonsil
CC and peripheral blood leukocytes, specially in plasmacytoid pre-
CC dendritic cells. Levels are much lower in monocytes and CD11c+
CC immature dendritic cells. Also detected in lung and liver.
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -!- SIMILARITY: CONTAINS 26 LEUCINE-RICH REPEATS (LRR).
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF259262; AAF72189.1; -
CC EMBL; AF259263; AAF72190.1; -
CC EMBL; AF245704; AAF78037.1; -
CC EMBL; AF246972; AAG01734.1; -
CC EMBL; AF246973; AAG01735.1; -
CC EMBL; AF246974; AAG01736.1; -
CC EMBL; AB045180; BAB19259.1; -
CC MIM; 605474; -
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003592; LRR_Out.
CC InterPro; IPR003591; LRR_tyr.
CC InterPro; IPR000157; TIR.
CC Pfam; PF00560; LRR; 13.
CC Pfam; PF01582; TIR; 1.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00370; LRR; 4.
CC SMART; SM00369; LRR_TVP; 2.
CC SMART; SM00255; TIR; 1.
CC PROSITE; PS50104; TIR; 1.
CC Receptor; immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein;
KW Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1032 TOLL-LIKE RECEPTOR 9.
FT DOMAIN 26 818 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 819 839 POTENTIAL.
FT DOMAIN 840 1032 CYTOPLASMIC (POTENTIAL).
FT REPEAT 62 85 LRR 1.
```

```
FT REPEAT 87 110 LRR 2.
FT REPEAT 122 147 LRR 3.
FT REPEAT 150 166 LRR 4.
FT REPEAT 167 190 LRR 5.
FT REPEAT 198 221 LRR 6.
FT REPEAT 223 242 LRR 7.
FT REPEAT 243 268 LRR 8.
FT REPEAT 283 306 LRR 9.
FT REPEAT 308 332 LRR 10.
FT REPEAT 333 356 LRR 11.
FT REPEAT 363 386 LRR 12.
FT REPEAT 390 413 LRR 13.
FT REPEAT 415 440 LRR 14.
FT REPEAT 470 494 LRR 15.
FT REPEAT 496 519 LRR 16.
FT REPEAT 520 543 LRR 17.
FT REPEAT 545 572 LRR 18.
FT REPEAT 574 598 LRR 19.
FT REPEAT 600 622 LRR 20.
FT REPEAT 627 650 LRR 21.
FT REPEAT 652 675 LRR 22.
FT REPEAT 676 699 LRR 23.
FT REPEAT 701 723 LRR 24.
FT REPEAT 724 747 LRR 25.
FT REPEAT 749 772 LRR 26.
FT DOMAIN 868 1016 TIR.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 694 694 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 731 731 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1 57 MISSING (IN ISOFORM 2).
FT VARSPLIC 1 1 M -> MPKWSGWSGWPATHTALPPQ (IN ISOFORM 3).
FT VARSPLIC 1 1 M -> MLYSSCKSLRLDSVEQDFHLEIAKK (IN ISOFORM 4).
FT VARSPLIC 1 16 MGFCRSALHPLSLVQ -> M (IN ISOFORM 5).
FT CONFLICT 530 530 H -> R (IN REF. 2; AAF78037).
FT CONFLICT 688 688 Q -> R (IN REF. 2; AAF78037).
SQ SEQUENCE 1032 AA; 115859 MW; 71280AA9680EDCE2 CRC64;

alignment_scores:
  Quality: 11.00 Length: 11
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x TLR9_HUMAN
Align seg 1/1 to: TLR9_HUMAN from: 1 to: 1032
2860 TTATGTCGAGGAAGGACTGGTTACCAAGG 2892
|||||
906 LeuCysLeuGluArgAspTrpLeuProGly 916
seq_name: SwissProt_40:TLR8_HUMAN
seq_documentation_block:
ID TLR8_HUMAN STANDARD; PRT; 1041 AA.
AC Q9NR97: Q9NYG9;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Toll-like receptor 8 precursor.
```

```
GN TLR8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20477806; PubMed=11022119;
RA Du X., Poltorak A., Wei Y., Beutler B.;
RT "Three novel mammalian Toll-like receptors: gene structure,
RT expression, and evolution.";
RL Eur. Cytokine Netw. 11:362-371(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=20477807; PubMed=11022120;
RA Chuang T.-H., Olevitch R.J.;
RT "Cloning and characterization of a sub-family of human Toll-like
RT receptors: hTLR7, hTLR8 and hTLR9.";
RL Eur. Cytokine Netw. 11:372-378(2000).
CC -!- FUNCTION: Participates in the innate immune response to microbial
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response (By
CC similarity).
CC -!- SUBUNIT: Binds MyD88 via their respective TIR domains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: Detected in brain, heart, lung, liver,
CC placenta, in monocytes, and at lower levels in CD11c+ immature
CC dendritic cells.
CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF246971; AAF64061.1; -.
CC EMBL; AF245703; AAF78036.1; -.
CC MIM; 300366;
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR_Cterm.
CC InterPro; IPR003592; LRR_Out.
CC InterPro; IPR003591; LRR_typ.
CC InterPro; IPR00157; TIR.
CC Pfam; PF00560; LRR; 16.
CC Pfam; PF01463; LRRCT; 1.
CC Pfam; PF01582; TIR; 1.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00370; LRR; 3.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00369; LRR_TYP; 3.
CC SMART; SM00255; TIR; 1.
CC PROSITE; PS50104; TIR; 1.
KW Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
KW SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1041 TOLL-LIKE RECEPTOR 8.
FT DOMAIN 27 827 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 828 848 POTENTIAL.
FT DOMAIN 849 1041 CYTOPLASMIC (POTENTIAL).
FT REPEAT 61 85 LRR 1.
FT REPEAT 87 112 LRR 2.
FT REPEAT 124 147 LRR 3.
FT REPEAT 169 198 LRR 4.
FT REPEAT 200 223 LRR 5.
FT REPEAT 225 244 LRR 6.
```



DR EMBL; AF138300; AAD44713.1; -  
DR EMBL; AF138301; AAF61437.1; -  
DR EMBL; AF138302; AAD44714.1; -  
DR EMBL; AF138303; AAF61438.1; -  
DR EMBL; AF138304; AAD44715.1; -  
DR PIR; A26476; NBHUC8.  
DR PIR; S05640; S05640.  
DR PIR; B28457; B28457.  
DR MIM; A45016; A45016.  
DR MIM; 125255; -  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_Out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00370; LRR; 3.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_Typ; 1.  
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
KW Repeat; Leucine-rich repeat; Signal; Alternative splicing;  
KW Polymorphism.  
FT SIGNAL 1 16 POTENTIAL.  
FT PROPEP 17 30  
FT CHAIN 31 359 BONE PROTEOGLYCAN II.  
FT REPEAT 77 98 LRR 1.  
FT REPEAT 99 122 LRR 2.  
FT REPEAT 123 145 LRR 3.  
FT REPEAT 146 167 LRR 4.  
FT REPEAT 168 193 LRR 5.  
FT REPEAT 194 217 LRR 6.  
FT REPEAT 218 238 LRR 7.  
FT REPEAT 239 262 LRR 8.  
FT REPEAT 263 285 LRR 9.  
FT REPEAT 286 308 LRR 10.  
FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN).  
FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT DISULFID 54 67 BY SIMILARITY.  
FT DISULFID 313 346 BY SIMILARITY.  
FT VARSPLIC 71 179 MISSING (IN ISOFORM B).  
FT VARSPLIC 73 219 MISSING (IN ISOFORM C).  
FT VARSPLIC 109 295 MISSING (IN ISOFORM D).  
FT VARSPLIC 72 75 LDKV -> CLPS (IN ISOFORM E).  
FT VARSPLIC 76 359 MISSING (IN ISOFORM E).  
FT VARIANT 273 273 E -> Q (IN DBSNP:1803344).  
FT CONFLICT 37 37 /FTIG-VAR\_011975.  
FT CONFLICT 45 45 G -> A (IN REF. 6).  
FT CONFLICT 45 45 D -> P (IN REF. 6).  
SQ SEQUENCE 359 AA; 39746 MW; FF5118871A1A52DD CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x PGS2\_HUMAN ..  
Align seg 1/1 to: PGS2\_HUMAN from: 1 to: 359  
508 ATACCGCAGGCGCTCCCGCCTAGCTTA 534  
|||||  
215 IleProGlnGlyLeuProPserLeu 223  
seq\_name: SwissProt\_40:PGS2\_BOVIN  
seq\_documentation\_block:  
ID PGS2\_BOVIN STANDARD; PRT; 360 AA.  
AC P21793;  
DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Bone proteoglycan II precursor (PG-S2) (Decorin).  
GN DCN.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88133946; PubMed=3435485;  
RA Day A.A., McQuillan C.I., Termino J.D., Young M.R.;  
RT "Molecular cloning and sequence analysis of the cDNA for small  
RT proteoglycan II of bovine bone.";  
RL Biochem. J. 248:801-805(1987).  
RN [2]  
RP SEQUENCE OF 31-54.  
RX MEDLINE=89123388; PubMed=2914936;  
RA Choi H.U., Johnson T.L., Pal S., Tang L.H., Rosenberg L., Neame P.J.;  
RT "Characterization of the dermatan sulfate proteoglycans, DS-PGI and  
RT DS-PGII, from bovine articular cartilage and skin isolated by octyl-  
RT sepharose chromatography.";  
RL J. Biol. Chem. 264:2876-2884(1989).  
CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE  
CC RATE OF FIBRIL FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-  
CC BETA.  
CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER  
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE  
CC TISSUE OF ORIGIN.  
CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS  
CC FAMILY.  
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
DR EMBL; Y00712; CAA68702.1; -  
DR PIR; S06280; S06280.  
DR PIR; B31430; B31430.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_Out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00370; LRR; 2.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_Typ; 2.  
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
KW Repeat; Leucine-rich repeat; Signal.  
FT SIGNAL 1 16 POTENTIAL.  
FT PROPEP 17 30  
FT CHAIN 31 360 BONE PROTEOGLYCAN II.  
FT REPEAT 78 99 LRR 1.  
FT REPEAT 100 123 LRR 2.  
FT REPEAT 124 146 LRR 3.  
FT REPEAT 147 168 LRR 4.  
FT REPEAT 169 194 LRR 5.  
FT REPEAT 195 218 LRR 6.  
FT REPEAT 219 239 LRR 7.  
FT REPEAT 240 263 LRR 8.  
FT REPEAT 264 286 LRR 9.  
FT REPEAT 287 309 LRR 10.  
FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 314 347 BY SIMILARITY.
SQ SEQUENCE 360 AA; 39837 MW; 71E84DA2D87552C0 CRC64;

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x PGS2_BOVIN ..
Align seg 1/1 to: PGS2_BOVIN from: 1 to: 360

508 ATACCGAGGCGCTCCGGCTAGCTTA 534
|||||
216 IleProGInGlyLeuProSerLeu 224

seq_name: SwissProt_40:PGS2_CANFA

seq_documentation_block:
ID PGS2_CANFA STANDARD; PRT; 360 AA.
AC Q29393;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone proteoglycan II precursor (PG-S2) (Decorin).
GN DCN OR DCNIC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Glant T.T.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 244-259 FROM N.A.
RA Venta P.J., Brouillette J.A., Yuzbasiyan-Gurkan V., Brewer G.J.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE
CC RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-
CC BETA (BY SIMILARITY).
CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE
CC TISSUE OF ORIGIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS
CC FAMILY.
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U83141; AAB51245.1; .
DR EMBL; L77684; AAA98062.1; .
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR0003592; LRR_out.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 1.
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;
Repeat; Leucine-rich repeat; Signal.
```

```
FT SIGNAL 1 16
FT PROPEP 17 30 POTENTIAL.
FT CHAIN 31 360 BY SIMILARITY.
FT REPEAT 78 99 BONE PROTEOGLYCAN II.
FT REPEAT 100 123 LRR 1.
FT REPEAT 124 146 LRR 2.
FT REPEAT 147 168 LRR 3.
FT REPEAT 169 194 LRR 4.
FT REPEAT 195 218 LRR 5.
FT REPEAT 219 239 LRR 6.
FT REPEAT 240 263 LRR 7.
FT REPEAT 264 286 LRR 8.
FT REPEAT 287 309 LRR 9.
FT CARBOHYD 34 34 O-LINKED (GLYCOSAMINOGLYCAN) (BY
FT CARBOHYD 212 212 SIMILARITY).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 314 347 BY SIMILARITY.
SQ SEQUENCE 360 AA; 39980 MW; 99BEE1A9C812906 CRC64;

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x PGS2_CANFA ..
Align seg 1/1 to: PGS2_CANFA from: 1 to: 360

508 ATACCGAGGCGCTCCGGCTAGCTTA 534
|||||
216 IleProGInGlyLeuProSerLeu 224

seq_name: SwissProt_40:PGS2_PIG

seq_documentation_block:
ID PGS2_PIG STANDARD; PRT; 360 AA.
AC Q9XSD9; Q9XSH4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone proteoglycan II precursor (PG-S2) (Decorin).
GN DCN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A. (LONG FORM).
RC STRAIN-YORKSHIRE;
RA Stephenson S., Schnoke M., Vesely I.;
RT "Cloning of the porcine decorin gene.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC STRAIN-YORKSHIRE; TISSUE-Aorta;
RA Stephenson S., Schnoke M., Vesely I.;
RT "Alternatively spliced version of the porcine decorin gene.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE
CC RATE OF FIBRILS FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-
CC BETA (BY SIMILARITY).
CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE
CC TISSUE OF ORIGIN.
CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS
CC FAMILY.
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AF125537; AAD23578.1; -;  
DR EMBL; AF140270; AAD33862.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_type.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00370; LRR; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
KW Repeat; Leucine-rich repeat; Signal; Alternative splicing.  
FT SIGNAL; 1 16  
FT PROPEP 17 30  
FT CHAIN 31 360 BONE PROTEOGLYCAN II.  
FT REPEAT 78 99  
FT REPEAT 100 123 LRR 1.  
FT REPEAT 124 146 LRR 2.  
FT REPEAT 147 168 LRR 3.  
FT REPEAT 169 194 LRR 4.  
FT REPEAT 195 218 LRR 5.  
FT REPEAT 219 239 LRR 6.  
FT REPEAT 240 263 LRR 7.  
FT REPEAT 264 286 LRR 8.  
FT REPEAT 287 309 LRR 9.  
FT REPEAT 309 340 LRR 10.  
FT CARBOHYD O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 55 68 BY SIMILARITY.  
FT DISULFID 314 347 POTENTIAL.  
FT VARSPIC 281 318 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 360 AA; 39899 MW; 8573DE8DBEA7509 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Smilarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x PGS2\_PIG

Align seg 1/1 to: PGS2\_PIG from: 1 to: 360

508 ATACCGAGGCGCTCCCGCTAGCTTA 534  
|||||

216 IleProGlnGlyLeuProSerLeu 224

seq\_name: SwissProt\_40:PGS2\_RABIT

seq\_documentation\_block:

ID PGS2\_RABIT STANDARD; PRT; 360 AA.  
AC Q28888; Q28608;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Bone proteoglycan II precursor (PG-S2) (Decorin).  
GN DCN.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;

RN SEQUENCE FROM N.A.  
RP TISSUE=Cornea;  
RX MEDLINE=9512319; PubMed=7822148;  
RA Zhan Q., Burrows R., Clintron C.;  
RT "Cloning and in situ hybridization of rabbit decorin in corneal  
RT tissues.";  
RL Invest. Ophthalmol. Vis. Sci. 36:206-215(1995).  
-----  
RN [2]  
RP SEQUENCE OF 38-358 FROM N.A.  
RC TISSUE=Cartilage;  
RA Hering T.M., Kollar J.;  
RT "The primary structure of rabbit chondrocyte decorin deduced from  
RT nucleotide sequence.";  
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: BINDS TO TYPE I AND TYPE II COLLAGEN AND AFFECTS THE  
CC RATE OF FIBRIL FORMATION. ALSO BINDS TO FIBRONECTIN AND TGF-  
CC BETA.  
CC -!- PTM: THE GLYCOSAMINOGLYCAN CHAIN ATTACHED TO DECORIN CAN BE EITHER  
CC CHONDROITIN SULFATE OR DERMATAN SULFATE DEPENDING UPON THE  
CC TISSUE OF ORIGIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SMALL INTERSTITIAL PROTEOGLYCANS  
CC FAMILY.  
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; S76584; AAB33083.1; -;  
DR EMBL; U03394; AAC04315.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_type.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01462; LRRNT; 1.  
DR SMART; SM00370; LRR; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 1.  
KW Glycoprotein; Connective tissue; Extracellular matrix; Proteoglycan;  
KW Repeat; Leucine-rich repeat; Signal.  
FT SIGNAL; 1 16  
FT PROPEP 17 30 BY SIMILARITY.  
FT CHAIN 31 360 BONE PROTEOGLYCAN II.  
FT REPEAT 78 99 LRR 1.  
FT REPEAT 100 123 LRR 2.  
FT REPEAT 124 146 LRR 3.  
FT REPEAT 147 168 LRR 4.  
FT REPEAT 169 194 LRR 5.  
FT REPEAT 195 218 LRR 6.  
FT REPEAT 219 239 LRR 7.  
FT REPEAT 240 263 LRR 8.  
FT REPEAT 264 286 LRR 9.  
FT REPEAT 287 309 LRR 10.  
FT CARBOHYD O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 360 AA; 39896 MW; 0B50C8756FE02369 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Smilarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x PGS2\_PIG

US-09-202-054-2 x PGS2\_RABIT ..  
Align seg 1/1 to: PGS2\_RABIT from: 1 to: 360

508 ATACCGAGCGCTCCGCGCTACGTTA 534  
|||||  
216 ILeProGlnGlyLeuProSerLeu 224

seq\_name: SwissProt\_40:FV1\_MOUSE

seq\_documentation\_block:

ID FV1\_MOUSE STANDARD; PRT: 459 AA.

AC P70213; P70214;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Friend virus susceptibility protein 1.

GN FV1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AKR/J, C3H, BALB/C, DBA/2, AND C57BL/6;

RX MEDLINE=96351076; PubMed=8752279;

RA Best S., le tissier P., Towers G., Stoye J.P.;

RT "Positional cloning of the mouse retrovirus restriction gene Fv1.";

RL Nature 382:826-829(1996).

CC -!- FUNCTION: CONTROLS REPLICATION OF THE MURINE LEUKEMIA VIRUS BY

CC INTERACTING WITH THE CAPSID PROTEIN CA AFTER ENTRY OF THE VIRUS

CC INTO THE CELL BUT BEFORE INTEGRATION AND FORMATION OF THE

CC PROVIRUS.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X97719; CAA66305.1; -.

DR EMBL; X97720; CAA66306.1; -.

DR MGD; MGI:95595; FV1.

ET VARIANT 358 358 E -> K (IN STRAINS AKR/J, C3H AND DBA/2).

ET VARIANT 399 399 R -> V (IN STRAINS AKR/J, C3H AND DBA/2).

ET VARIANT 438 459 GLTSGVSGVLSLSPWKRQNS -> TKL (IN

ET SEQUENCE 459 AA; 51996 MW; B112FD4645EFA489 CRC64;

alignment\_scores:

Quality: 9.00 Length: 9

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x FV1\_MOUSE ..

Align seg 1/1 to: FV1\_MOUSE from: 1 to: 459

3101 GGCTGTGGGAGTCTGTGCTTGTGCT 3127

|||||

443 GlySerValGlyValLeuSerLeuSer 451

seq\_name: SwissProt\_40:GPBA\_HUMAN

seq\_documentation\_block:

ID GPBA\_HUMAN STANDARD; PRT: 626 AA.

AC P07359;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Platelet glycoprotein Ib alpha chain precursor (GP-IB alpha) (GP1BA)  
DE (CD42B-alpha) (CD42B) [Contains: Glycocalicin].  
GN GP1BA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE=87289655; PubMed=3303030;

RA Lopez J.A., Chung D.W., Fujikawa K., Hagen F.S., Papayannopoulou T.,

RA Roth G.J.;

RT "Cloning of the alpha chain of human platelet glycoprotein Ib: a

RT transmembrane protein with homology to leucine-rich alpha 2-

RT glycoprotein.";

RL Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=89025874; PubMed=2845978;

RA Wenger R.H., Kieffer N., Wicki A.N., Clemetson K.J.;

RT "Structure of the human blood platelet membrane glycoprotein Ib alpha

RT gene.";

RL Biochem. Biophys. Res. Commun. 156:389-395(1988).

RN [3]

RP SEQUENCE OF 17-315.

RX MEDLINE=87289654; PubMed=3497398;

RA Tisani K., Takio K., Handa M., Ruggeri Z.M.;

RT "Amino acid sequence of the von Willebrand factor-binding domain of

RT platelet membrane glycoprotein Ib.";

RL Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614(1987).

RN [4]

RP DISULFIDE BONDS.

RX MEDLINE=91301149; PubMed=2070794;

RA Hess D., Schaller J., Rickli E.E., Clemetson K.J.;

RT "Identification of the disulphide bonds in human platelet

RT glycoocalicin.";

RL Eur. J. Biochem. 199:389-393(1991).

RN [5]

RP VARIANT SIBA

RX MEDLINE=92265982; PubMed=1586750;

RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,

RA Ruggeri Z.M.;

RT "Genetic and structural characterization of an amino acid dimorphism

RT in glycoprotein Ib alpha involved in platelet transfusion

RT refractoriness.";

RL Blood 79:3086-3090(1992).

RN [6]

RP VARIANT BSS PHE-73.

RX MEDLINE=92110577; PubMed=1730088;

RA Miller J.L., Lyle V.A., Cunningham D.;

RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein

RT Ib alpha leucine tandem repeat occurring in patients with an

RT autosomal dominant variant of Bernard-Soulier disease.";

RL Blood 79:439-446(1992).

RN [7]

RP VARIANT BSS VAL-172.

RX MEDLINE=93388851; PubMed=7690774;

RA Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,

RA de Marco L., Ruggeri Z.M.;

RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib

RT alpha resulting in the Bernard-Soulier syndrome.";

RL J. Clin. Invest. 92:1213-1220(1993).

RN [8]

RP VARIANT BSS SER-225.

RX MEDLINE=95118882; PubMed=7819107;

RA Simsek S., Noris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,

RA Ribera A., Gallardo D.;

RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha

RT gene is associated with Bernard-Soulier syndrome.";

RL Br. J. Haematol. 88:839-844(1994).

RN [9]

RP VARIANT PSEUDO-VMD VAL-249.



RX MEDLINE=91271273; PubMed=2052556;  
RA Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;  
RT "Mutation in the gene encoding the alpha chain of platelet  
RL glycoprotein Ib in platelet-type von Willebrand disease.";   
RN Proc. Natl. Acad. Sci. U.S.A. 88:4761-4765(1991).  
RP [10]  
RQ VARIANT PSEUDO-VWD VAL-249.  
RX MEDLINE=93253059; PubMed=8486780;  
RA Murata M., Russell S.R., Ruggeri Z.M., Ware J.;  
RT "Expression of the phenotypic abnormality of platelet-type von  
RL Willebrand disease in a recombinant glycoprotein Ib alpha fragment.";   
RN J. Clin. Invest. 91:2133-2137(1993).  
RP [11]  
RQ VARIANT PSEUDO-VWD VAL-255.  
RX MEDLINE=93214031; PubMed=8384898;  
RA Russell S.D., Roth G.J.;  
RT "Pseudo-von Willebrand disease: a mutation in the platelet  
RL glycoprotein Ib alpha gene associated with a hyperactive surface  
RN receptor.";   
RP [12]  
RQ VARIANT BSS LEU-195 DEL.  
RX MEDLINE=95178321; PubMed=7873390;  
RA de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,  
RT Chevalier J., Gachet C., Briquel M.-E., Cazenave J.-P.;  
RL "A three-base deletion removing a leucine residue in a leucine-rich  
RN repeat of platelet glycoprotein Ib alpha associated with a variant of  
RP Bernard-Soulier syndrome (Nancy I).";   
RQ Br. J. Haematol. 89:386-396(1995).  
RX CC -1- FUNCTION: GP-IB, A SURFACE MEMBRANE PROTEIN OF PLATELETS,  
CC PARTICIPATES IN THE FORMATION OF PLATELET PLUGS BY BINDING TO VON  
CC WILLEBRAND FACTOR, WHICH IS ALREADY BOUND TO THE SUBENDOTHELIUM.  
CC -1- SUBUNIT: GP-IB ALPHA AND BETA ARE DISULFIDE LINKED. GP-IX IS  
CC COMPLEXED WITH THE GP-IB HETERODIMER VIA A NON COVALENT LINKAGE.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- PTM: GLYCOCALICIN, WHICH IS APPROXIMATELY COEXTENSIVE WITH THE  
CC EXTRACELLULAR PART OF THE MOLECULE, IS CLEAVED OFF BY CALPAIN  
CC DURING PLATELET LYSIS.  
CC -1- POLYMORPHISM: POSITION 161 IS ASSOCIATED WITH PLATELET-SPECIFIC  
CC ALLOANTIGEN SIBA. SIBA(-) HAS THR-161 AND SINA(+) HAS MET-161.  
CC SIBA IS INVOLVED IN NEONATAL ALLOIMMUNE THROMBOCYTOPENIA (NATP).  
CC -1- DISEASE: DEFECTS IN GP-IB ARE ONE OF THE CAUSES OF BERNARD-SOULIER  
CC SYNDROME (BSS). BSS PATIENTS HAVE UNUSUALLY LARGE PLATELETS AND  
CC HAVE A CLINICAL BLEEDING TENDENCY.  
CC -1- DISEASE: DEFECTS IN GP-IB ARE ONE OF THE CAUSES OF VON WILLEBRAND  
CC DISEASE (VWD) KNOWN AS PLATELET-TYPE VON WILLEBRAND DISEASE OR  
CC PSEUDO-VON WILLEBRAND DISEASE. THIS AUTOSOMAL DOMINANT BLEEDING  
CC DISORDER IS CAUSED BY AN INCREASED AFFINITY OF GP-IB FOR SOLUBLE  
CC VWF RESULTING IN IMPAIRED HEMOSTATIC FUNCTION DUE TO THE REMOVAL  
CC OF VWF FROM THE CIRCULATION.  
CC -1- MISCELLANEOUS: PLATELET ACTIVATION APPARENTLY INVOLVES DISRUPTION  
CC OF THE MACROMOLECULAR COMPLEX OF GP-IB WITH THE PLATELET  
CC ACTIN-BINDING PROTEIN.  
CC GLYCOPROTEIN IX (GP-IX) AND DISSOCIATION OF GP-IB FROM THE  
CC -1- MISCELLANEOUS: BINDING SITES FOR VON WILLEBRAND FACTOR AND  
CC THROMBIN (THE LATTER SITE WITH UNKNOWN FUNCTION) ARE IN THE  
CC AMINO-TERMINAL PART OF THE MOLECULE.  
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02940; AAA52595.1; -;  
DR EMBL; M22403; AAA52596.1; -;  
DR PIR; A27075; NBHUIA. -;  
DR GlycoSiteDB; P07359; -;  
DR MIM; 231200; -;  
DR MIM; 177820; -;

DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_typ.  
DR Pfam; PF00560; LRR; 6.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 3.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
KW Platelet; Transmembrane; Glycoprotein; Hemostasis; Blood coagulation;  
KW Repeat; Leucine-rich repeat; Signal; Cell adhesion; Disease mutation;  
KW Polymorphism; von Willebrand disease; Bernard Soulier syndrome.  
FT SIGNAL 1 16  
FT CHAIN 17 626 PLATELET GLYCOPROTEIN IB ALPHA CHAIN.  
FT CHAIN 17 ? GLYCOCALICIN.  
FT DOMAIN 17 505 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 506 526 POTENTIAL.  
FT DOMAIN 527 626 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 70 92 LRR 1.  
FT REPEAT 93 117 LRR 2.  
FT REPEAT 119 138 LRR 3.  
FT REPEAT 139 162 LRR 4.  
FT REPEAT 164 186 LRR 5.  
FT REPEAT 188 210 LRR 6.  
FT REPEAT 379 386 THR/PRO-RICH.  
FT REPEAT 387 395 THR/PRO-RICH.  
FT REPEAT 400 408 THR/PRO-RICH.  
FT REPEAT 409 417 THR/PRO-RICH.  
FT REPEAT 422 430 THR/PRO-RICH.  
FT DISULFID 20 33  
FT DISULFID 225 264  
FT DISULFID 227 280  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .).  
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .).  
FT CARBOHYD 308 308 O-LINKED.  
FT VARIANT 72 72 R -> H (IN DBSNP:6068).  
FT VARIANT 73 73 L -> F (IN BSS).  
FT VARIANT 161 161 /FTID=VAR\_005256.  
FT VARIANT 172 172 /FTID=VAR\_005257.  
FT VARIANT 172 172 A -> V (IN BSS).  
FT VARIANT 195 195 /FTID=VAR\_005258.  
FT VARIANT 195 195 MISSING (IN BSS).  
FT VARIANT 225 225 /FTID=VAR\_005259.  
FT VARIANT 225 225 C -> S (IN BSS).  
alignment\_scores: Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x GPBA\_HUMAN ..  
Align seg 1/1 to: GPBA\_HUMAN from: 1 to: 626  
2191 ACTTTGGAGCTCAGCCACCAACCACTG 2217  
|||||  
97 ThrLeuAspLeuSerHisAsnGlnLeu 105  
seq\_name: SwissProt\_40:TLR4\_PAPAN  
seq\_documentation\_block:  
ID TLR4\_PAPAN STANDARD; PRT; 826 AA.  
AC Q9TSP2;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Toll-like receptor 4 precursor.  
GN TLR4.  
OS Papio anubis (Olive baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Papio.  
OX NCBI\_TaxID=9555;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PubMed=11104518;  
RA Smirnova I., Poltorak A., Chan E.K.L., McBride C., Beutler B.;  
RT "Phylogenetic variation and polymorphism at the Toll-like receptor 4  
locus (TLR4).";  
RL (In) Genome Biol. 1:RESEARCH002.1-2.10(2000).  
CC -|- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate  
immune response to bacterial lipopolysaccharide (LPS). Acts via  
MyD88, TRAP and TRAF6, leading to NF-kappa-B activation, cytokine  
secretion and the inflammatory response (By similarity).  
CC -|- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a  
multi-protein complex containing at least CD14, MD-2 and TLR4.  
CC Binds MD-2 via the extracellular domain. Binds MyD88 and TRAP via  
their respective TIR domains (By similarity).  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -|- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -|- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -|- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF180964; AAF07059.1; .  
DR EMBL; AF180962; AAF07059.1; JOINED.  
DR EMBL; AF180963; AAF07059.1; JOINED.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_Typ.  
DR InterPro; IPR00157; TIR.  
DR Pfam; PF00560; LRR; 9.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01582; TIR; 1.  
DR PRINTS; PRO00019; LEURICHRPT.  
DR SMART; SM00370; LRR; 1.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00369; LRR\_TYP; 2.  
DR SMART; SM00255; TIR; 1.  
DR PROSITE; PS50104; TIR; 1.  
KW Receptor; Immune response; Inflammatory response; Signal;  
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.  
FT SIGNAL 1 23  
FT CHAIN 24 826  
FT DOMAIN 24 631  
FT DOMAIN 632 652  
FT DOMAIN 653 826  
FT REPEAT 53 76  
FT REPEAT 77 100  
FT REPEAT 101 124  
FT REPEAT 128 149  
FT REPEAT 150 173  
FT REPEAT 174 197  
FT REPEAT 203 225  
FT REPEAT 228 252  
FT REPEAT 277 303  
FT REPEAT 327 350  
FT REPEAT 351 372  
FT REPEAT 373 398  
FT REPEAT 400 421  
FT REPEAT 422 445

FT REPEAT 447 469  
FT REPEAT 470 494  
FT REPEAT 495 518  
FT REPEAT 520 541  
FT REPEAT 543 569  
FT REPEAT 571 592  
FT DOMAIN 672 818  
FT CARBOHYD 35 35  
FT CARBOHYD 173 173  
FT CARBOHYD 205 205  
FT CARBOHYD 282 282  
FT CARBOHYD 309 309  
FT CARBOHYD 497 497  
FT CARBOHYD 526 526  
FT CARBOHYD 575 575  
FT CARBOHYD 624 624  
FT CARBOHYD 630 630  
SQ SEQUENCE 826 AA; 94678 MW; 422777318E5F1769 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x TLR4\_PAPAN ..  
Align seg 1/1 to: TLR4\_PAPAN from: 1 to: 826  
2338 CTGGATCTCAGTCAATAAATAATCCAG 2364  
180 LeuAspLeuSerSerAsnLysIleGln 188  
|||||  
seq\_name: SwissProt\_40:TLR4\_HUMAN  
seq\_documentation\_block:  
ID TLR4\_HUMAN STANDARD; PRT; 839 AA.  
AC O00206; Q9UK78; Q9UM57;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Toll-like receptor 4 precursor (hToll).  
GN TLR4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Spleen;  
RX MEDLINE=97379437; PubMed=92377759;  
RA Medzhitov R., Preston-Hurlburt P., Janeway C.A. Jr.;  
RT "A human homologue of the Drosophila Toll protein signals activation  
of adaptive immunity.";  
RL Nature 388:394-397(1997).  
RN [2]  
RP SEQUENCE OF 41-839 FROM N.A.  
RC TISSUE=Lung, Placenta, and Fetal liver;  
RX MEDLINE=98118556; PubMed=9435236;  
RA Rock F.L., Hardiman G., Timans J.C., Kastelein R.A., Bazan J.F.;  
RT "A family of human receptors structurally related to Drosophila  
Toll.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANTS GLY-299 AND ILE-399.  
RX PubMed=11104518;  
RA Smirnova I., Poltorak A., Chan E.K.L., McBride C., Beutler B.;  
RT "Phylogenetic variation and polymorphism at the Toll-like receptor 4  
locus (TLR4).";  
RL (In) Genome Biol. 1:RESEARCH002.1-2.10(2000).  
RN [4]  
RP SEQUENCE FROM N.A., AND VARIANTS GLY-299 AND ILE-399.

RX PubMed-10835634;  
RA Arbour N.C., Lorenz E., Schutte B.C., Zabner J., Kline J.N., Jones M.,  
RA Frees K., Watt J.L., Schwartz D.A.;  
RT "TLR4 mutations are associated with endotoxin hyporesponsiveness in  
humans";  
RL Nat. Genet. 25:187-191(2000).  
RN [5]  
RP MUTAGENESIS OF GLU-697; ARG-710; ASP-711 AND PRO-714.  
RX PubMed-11081518;  
RA Xu Y., Tao X., Shen B., Horng T., Medzhitov R., Manley J.L., Tong L.;  
RT "Structural basis for signal transduction by the Toll/interleukin-1  
receptor domains";  
RL Nature 408:111-115(2000).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND MUTAGENESIS OF ASN-526 AND ASN-575.  
RX PubMed-11706042;  
RA da Silva Correia J., Ulevitch R.J.;  
RT "MD-2 and TLR4 N-linked glycosylations are important for a functional  
lipopolysaccharide receptor";  
RL J. Biol. Chem. 277:1845-1854(2002).  
CC -|- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate  
immune response to bacterial lipopolysaccharide (LPS). Acts via  
MyD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine  
secretion and the inflammatory response.  
CC -|- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a  
multi-protein complex containing at least CD14, MD-2 and TLR4.  
CC Binds MD-2 via the extracellular domain. Binds MyD88 and TIRAP via  
their respective TIR domains.  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -|- TISSUE SPECIFICITY: Highly expressed in placenta, spleen and  
peripheral blood leukocytes. Detected in monocytes, macrophages,  
dendritic cells and several types of T-cells.  
CC -|- PM: N-glycosylated. Glycosylation of Asn-526 and Asn-575 seems to  
be necessary for the expression of TLR4 on the cell surface and  
the LPS-response. Likewise, mutants lacking two or more of the  
other N-glycosylation sites were deficient in interaction with  
LPS.  
CC -|- POLYMORPHISM: Allele B (Gly-299, Ile-399) is associated with a  
blunted response to inhaled LPS.  
CC -|- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.  
CC -|- SIMILARITY: CONTAINS 1 TIR DOMAIN.  
CC -|- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U93091; AAC80227.1; -;  
DR EMBL: U88880; AAC34135.1; -;  
DR EMBL: AF177765; AAF05316.1; -;  
DR EMBL: AF177766; AAF07823.1; -;  
DR EMBL: AF172171; AAF89753.1; -;  
DR EMBL: AF172169; AAF89753.1; JOINED.  
DR EMBL: AF172170; AAF89753.1; JOINED.  
DR MIM: 603030; -;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR000483; LRR\_Cterm.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR003591; LRR\_type.  
DR InterPro: IPR000157; TIR.  
DR Pfam: PF00560; LRR; 9.  
DR Pfam: PF01463; LRRCT; 1.  
DR Pfam: PF01582; TIR; 1.  
DR PRINTS: PR00019; LEURICHRPT.  
DR SMART: SM00370; LRR; 2.  
DR SMART: SM00082; LRRCT; 1.  
DR SMART: SM00369; LRR\_TYP; 2.  
DR SMART: SM00255; TIR; 1.  
DR PROSITE: PS50104; TIR; 1.

KW Receptor; Immune response; Inflammatory response; Signal;  
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein;  
KW Polymorphism.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 839 TOLL-LIKE RECEPTOR 4.  
FT DOMAIN 24 631 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 632 652 POTENTIAL.  
FT DOMAIN 653 839 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 52 76 LRR 1.  
FT REPEAT 77 100 LRR 2.  
FT REPEAT 101 124 LRR 3.  
FT REPEAT 128 149 LRR 4.  
FT REPEAT 150 173 LRR 5.  
FT REPEAT 174 197 LRR 6.  
FT REPEAT 203 225 LRR 7.  
FT REPEAT 228 252 LRR 8.  
FT REPEAT 277 303 LRR 9.  
FT REPEAT 307 330 LRR 10.  
FT REPEAT 332 350 LRR 11.  
FT REPEAT 351 372 LRR 12.  
FT REPEAT 373 398 LRR 13.  
FT REPEAT 400 421 LRR 14.  
FT REPEAT 422 445 LRR 15.  
FT REPEAT 447 469 LRR 16.  
FT REPEAT 470 494 LRR 17.  
FT REPEAT 495 518 LRR 18.  
FT REPEAT 520 541 LRR 19.  
FT REPEAT 543 566 LRR 20.  
FT REPEAT 568 592 LRR 21.  
FT DOMAIN 672 818 TIR.  
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .).  
FT VARIANT 299 299 D -> G (IN ALLELE B; REDUCED LPS-  
RESPONSE).  
FT VARIANT 399 399 /FTID=VAR\_012739.  
FT VARIANT 526 526 T -> I (IN ALLELE B; REDUCED LPS-  
RESPONSE).  
FT MUTAGEN 526 526 /FTID=VAR\_012740.  
FT MUTAGEN 575 575 N->A: ABOLISHES LPS-RESPONSE AND PREVENTS  
THE CELL SURFACE EXPRESSION.  
FT MUTAGEN 697 697 N->A: ABOLISHES LPS-RESPONSE AND PREVENTS  
THE CELL SURFACE EXPRESSION.  
FT MUTAGEN 710 710 E->R: ABOLISHES LPS-RESPONSE.  
FT MUTAGEN 711 710 R->E: ABOLISHES LPS-RESPONSE.  
FT MUTAGEN 714 711 D->K: ABOLISHES LPS-RESPONSE.  
FT MUTAGEN 714 711 P->H,R,E: ABOLISHES MYD88-BINDING AND  
LPS-RESPONSE.  
SQ SEQUENCE 839 AA; 95679 MW; 92C48F55821133E8 CRC64;  
  
alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x TLR4\_HUMAN ..  
Align seg 1/1 to: TLR4\_HUMAN from: 1 to: 839  
2338 CTGGATCTCAGCTCAATAATAATCCAG 2364  
|||||  
180 LeuaspLeuSerAsnlylleGln 188  
seq\_name: SwissProt\_40:TLR4\_PANPA



RL Proc. Natl. Acad. Sci. U.S.A. 90:11386-11390(1993).  
CC !- FUNCTION: MAY PLAY A KEY ROLE IN EMBRYONIC CELLULARIZATION BY  
CC INTERACTING WITH BOTH THE CYTOSKELETON AND OTHER CELLULAR  
CC COMPONENTS.  
CC !- TISSUE SPECIFICITY: STRONGEST EXPRESSION IN SKELETAL MUSCLE WITH  
CC HIGH EXPRESSION ALSO IN THE HEART AND LUNG.  
CC !- DISEASE: THIS PROTEIN IS ABSENT IN SMITH-MAGENIS SYNDROME (SMS), A  
CC RELATIVELY COMMON MICRODELETION SYNDROME INVOLVING DEVELOPMENTAL  
CC ABNORMALITIES AND MENTAL RETARDATION.  
CC !- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).  
CC !- SIMILARITY: CONTAINS 5 GELSOLIN-LIKE REPEATS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U01184; AAC02796.1; .  
DR EMBL; U01184; AAC03568.1; .  
DR HSSP; P02640; 2VIL.  
DR MIM; 600362; .  
DR InterPro; IPR001974; Gelsolin.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR003592; LRR.out.  
DR Pfam; PF00626; Gelsolin; 5.  
DR Pfam; PF00560; LRR; 11.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00262; GEL; 6.  
DR SMART; SM00370; LRR; 7.  
KW Developmental protein; Repeat; Leucine-rich repeat.  
FT REPEAT 7 32 LRR 1.  
FT REPEAT 32 55 LRR 2.  
FT REPEAT 56 78 LRR 3.  
FT REPEAT 80 103 LRR 4.  
FT REPEAT 104 126 LRR 5.  
FT REPEAT 127 149 LRR 6.  
FT REPEAT 150 173 LRR 7.  
FT REPEAT 175 196 LRR 8.  
FT REPEAT 197 222 LRR 9.  
FT REPEAT 223 245 LRR 10.  
FT REPEAT 247 268 LRR 11.  
FT REPEAT 269 291 LRR 12.  
FT REPEAT 293 316 LRR 13.  
FT REPEAT 318 339 LRR 14.  
FT REPEAT 340 363 LRR 15.  
FT REPEAT 501 559 GELSOLIN-LIKE 1.  
FT REPEAT 640 670 GELSOLIN-LIKE 2.  
FT REPEAT 755 798 GELSOLIN-LIKE 3.  
FT REPEAT 1088 1115 GELSOLIN-LIKE 4.  
FT REPEAT 1176 1218 GELSOLIN-LIKE 5.  
SQ SEQUENCE 1269 AA; 144750 MW; 29AC7C07738B7B47 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x FLIH\_HUMAN ..

Align seg 1/1 to: FLIH\_HUMAN from: 1 to: 1269

2194 TTGACCTCAGCCACACCAACTGACC 2220  
|||||  
109 LeuAspLeuSerHisAsnGlnLeuThr 117

seq name: SwissProt\_40:CHAO\_DROME

seq\_documentation\_block:

ID CHAO\_DROME STANDARD; PRT; 1315 AA.  
AC P12024; Q9VA01;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Chaptin precursor (photoreceptor cell-specific membrane protein).  
GN CHP OR CHT OR CGI744.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=88135762; PubMed=3124963;  
RA Renke R., Krantz D.E., Yen D., Zipursky S.L.;  
RT "Chaptin, a cell surface glycoprotein required for Drosophila  
RT photoreceptor cell morphogenesis, contains a repeat motif found in  
RT yeast and human.";  
RL Cell 52:291-301(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.V., Benos P.V., Beremand B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck C.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [3]  
RP SEQUENCE OF 30-50 FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=85166231; PubMed=3920657;  
RA Zipursky S.L., Venkatesh T.R., Benzer S.;  
RT "From monoclonal antibody to gene for a neuron-specific glycoprotein  
RT in Drosophila";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:1855-1859(1985).  
RN [4]  
RP SEQUENCE OF 30-50, AND TISSUE SPECIFICITY.

RC STRAIN-CANTON-S; TISSUE-Head;  
 RA MEDLINE=84106810; PubMed=6420071;  
 RT Zipursky S.L., Venkatesh T.R., Teplov D.B., Benzer S.;  
 RT "Neuronal development in the Drosophila retina: monoclonal antibodies  
 as molecular probes.";  
 RL Cell 36:15-26(1984).  
 CC -!- FUNCTION: REQUIRED FOR PHOTORECEPTOR CELL MORPHOGENESIS. MEDIATES  
 CC HOMOPHILIC CELLULAR ADHESION.  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR SURFACE OF R-CELL PLASMA  
 CC MEMBRANE.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PHOTORECEPTOR CELLS AND THEIR  
 CC AXONS IN THE ADULT RETINA, THE OCELLUS AND LARVAL PHOTORECEPTOR  
 CC ORGAN.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED 24 HOURS AFTER INITIATION OF  
 CC PHOTORECEPTOR CELL DIFFERENTIATION, PERSISTS THROUGH DEVELOPMENT  
 CC TO ADULTHOOD.  
 CC -!- SIMILARITY: BELONGS TO THE CHAOTIN FAMILY.  
 CC -!- SIMILARITY: CONTAINS 38 LEUCINE-RICH REPEATS (LRR).  
 CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
 CC FRAMESHIFT IN POSITION 1123.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M19017; AAA28425.1; ALT-FRAME.  
 CC EMBL; M19008; AAA28425.1; JOINED.  
 CC EMBL; M19009; AAA28425.1; JOINED.  
 CC EMBL; M19010; AAA28425.1; JOINED.  
 CC EMBL; M19011; AAA28425.1; JOINED.  
 CC EMBL; M19012; AAA28425.1; JOINED.  
 CC EMBL; M19013; AAA28425.1; JOINED.  
 CC EMBL; M19014; AAA28425.1; JOINED.  
 CC EMBL; M19016; AAA28425.1; JOINED.  
 CC EMBL; AE003777; AAF57127.1; --  
 CC EMBL; K03274; AAA280851.1; --  
 CC PIR; A29944; A29944.  
 CC FlyBase; FBgn0000313; chp.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR003592; LRR\_out.  
 CC InterPro; IPR003591; LRR\_typ.  
 CC Pfam; PF00560; LRR; 29  
 CC PRINTS; PR00019; LEURICHRPT.  
 CC SMART; SM00370; LRR; 9.  
 CC SMART; SM00369; LRR\_TYP; 27.  
 CC Glycoprotein; Membrane; Signal; Repeat; Leucine-rich repeat; Vision.  
 KW SIGNAL 1 29  
 FT CHAIN 30 1315 CHAOTIN.  
 FT REPEAT 101 124 LRR 1.  
 FT REPEAT 126 149 LRR 2.  
 FT REPEAT 150 173 LRR 3.  
 FT REPEAT 175 198 LRR 4.  
 FT REPEAT 199 222 LRR 5.  
 FT REPEAT 224 247 LRR 6.  
 FT REPEAT 249 272 LRR 7.  
 FT REPEAT 277 300 LRR 8.  
 FT REPEAT 302 324 LRR 9.  
 FT REPEAT 326 347 LRR 10.  
 FT REPEAT 349 372 LRR 11.  
 FT REPEAT 374 397 LRR 12.  
 FT REPEAT 451 474 LRR 13.  
 FT REPEAT 475 498 LRR 14.  
 FT REPEAT 525 548 LRR 15.  
 FT REPEAT 550 572 LRR 16.  
 FT REPEAT 575 598 LRR 17.  
 FT REPEAT 599 622 LRR 18.  
 FT REPEAT 624 646 LRR 19.  
 FT REPEAT 648 670 LRR 20.  
 FT REPEAT 674 697 LRR 21.

FT REPEAT 706 729 LRR 22.  
 FT REPEAT 731 754 LRR 23.  
 FT REPEAT 755 778 LRR 24.  
 FT REPEAT 779 802 LRR 25.  
 FT REPEAT 803 826 LRR 26.  
 FT REPEAT 828 849 LRR 27.  
 FT REPEAT 852 875 LRR 28.  
 FT REPEAT 877 900 LRR 29.  
 FT REPEAT 902 924 LRR 30.  
 FT REPEAT 926 947 LRR 31.  
 FT REPEAT 948 970 LRR 32.  
 FT REPEAT 971 994 LRR 33.  
 FT REPEAT 995 1017 LRR 34.  
 FT REPEAT 1043 1068 LRR 35.  
 FT REPEAT 1117 1140 LRR 36.  
 FT REPEAT 1142 1165 LRR 37.  
 FT REPEAT 1169 1193 LRR 38.  
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 680 680 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 746 746 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 936 936 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 970 970 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1012 1012 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1122 1122 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1152 1152 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1171 1171 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 12 12 A -> V (IN REF. 1).  
 FT CONFLICT 44 44 C -> H (IN REF. 4).  
 FT CONFLICT 50 50 C -> H (IN REF. 4).  
 FT CONFLICT 937 937 I -> V (IN REF. 1).  
 SQ SEQUENCE 1315 AA; 151982 MW; D2D89A64EB4FCE5 CRC64;  
 alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-202-054-2 x CHAO\_DROME ..  
 Align seg 1/1 to: CHAO\_DROME from: 1 to: 1315  
 835 TTACAATTCCTGACCTAAGTGAAT 861  
 (|||||)|||||  
 352 LeuGlnIleLeuAspLeuSerGlyAsn 360  
 seq\_name: SwissProt\_40:R115\_YEAST  
 seq\_documentation\_block:  
 ID R115\_YEAST STANDARD; PRT; 1770 AA.  
 AC P43565;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Serine/threonine-protein kinase RIM15 (EC 2.7.1.-).  
 GN RIM15 OR TAK1 OR YFL033C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=95400292; PubMed=7670463;

```
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [2].
RN
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97265402; PubMed=9111339;
RA Vidan S., Mitchell A.P.;
RT "Stimulation of yeast meiotic gene expression by the
RT glucose-repressible protein kinase Rim15p.";
RL Mol. Cell. Biol. 17:2688-2697(1997).
RN [3].
RN
RP SEQUENCE FROM N.A.
RA Reinders A., Buerckert N., Boller T., Wiemken A., de Virgilio C.;
RT "Saccharomyces cerevisiae CAMP-dependent protein kinase controls entry
RT into stationary phase through the Rim15p protein kinase.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE ACTIVATION OF A MEIOTIC GENES ACTIVATION
CC PATHWAY.
CC
CC -!- PTM: AUTOPHOSPHORYLATED.
CC
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
CC -!- STRONG, WITH S. POMBE CEK1.
CC
CC -!- SIMILARITY: CONTAINS 1 RESPONSE REGULATORY DOMAIN.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50617; BAA09206.1; -.
DR EMBL; U83459; AAB64088.1; -.
DR EMBL; AJ001030; CAA04486.1; -.
DR HSPF; Q16539; IWFC.
DR SGB; S0001861; RIM15.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 2.
DR Pfam; PF00072; response_reg; 1.
DR SMART; SM00448; REC; 1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation; Meiosis.
FT DOMAIN 794..1254 PROTEIN KINASE.
FT DOMAIN 1636..1750 RESPONSE REGULATORY.
FT DOMAIN 343..358 POLY-ASN.
FT DOMAIN 620..624 POLY-SER.
FT NP_BIND 800..808 ATP (BY SIMILARITY).
FT BINDING 823..823 ATP (BY SIMILARITY).
FT ACT_SITE 918..918 BY SIMILARITY.
FT DOMAIN 975..980 POLY-ASN.
FT DOMAIN 1213..1218 POLY-GLU.
FT DOMAIN 1386..1391 POLY-THR.
SQ SEQUENCE 1770 AA; 196530 MW; DC1064825000FAFF CRC64;
```

```
alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x MTRG_METTH
```

```
US-09-202-054-2 x R115_YEAST
Align seg 1/1 to: R115_YEAST from: 1 to: 1770

2107 CCNAATCTAAAGAAATCTCTTTGGCC 2133
1397 ProAsnLeuLysAsnLeuSerLeuAla 1405

seq_name: SwissProt_40:MTRG_METTH

seq_documentation_block:
ID MTRG_METTH STANDARD; PRT: 85 AA.
AC OT27225;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tetrahydromethanopterin S-methyltransferase subunit G (EC 2.1.1.86)
DE (N5-methyltetrahydromethanopterin--coenzyme M methyltransferase
DE subunit G).
DE GN MTRG OR MTH1157.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McQuigall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delah: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN
CC METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND N5-METHYL-
CC TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-
CC TETRAHYDROMETHANOPTERIN. THIS IS A ENERGY-CONSERVING, SODIUM-ION
CC TRANSLOCATING STEP.
CC -!- CATALYTIC ACTIVITY: 5-methyl-5,6,7,8-tetrahydromethanopterin + 2-
CC mercaptoethanesulfonate = 5,6,7,8-tetrahydromethanopterin + 2-
CC (methylthio)ethanesulfonate.
CC -!- SUBUNIT: COMPOSED OF 8 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000885; AAB85646.1; -.
DR Transferrase; Methyltransferase; Transmembrane; Methanogenesis;
KW Complete proteome.
FT INIT_MET 0 BY SIMILARITY.
FT TRANSMEM 51 71 POTENTIAL.
SQ SEQUENCE 85 AA; 9379 MW; C0BB9784DE54AAB1 CRC64;
```

```
alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x MTRG_METTH
```

Align seg 1/1 to: MTRG\_METTM from: 1 to: 85

3078 GGACTTCTGAAGGCTTCTCAAG 3055

|||||

73 GlyLeuLeuLysGlyLeuLeuLys 80

seq\_name: SwissProt\_40:MTRG\_METTM

seq\_documentation\_block:

ID MTRG\_METTM STANDARD; PRT; 85 AA.

AC Q50774;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Tetrahydromethanopterin S-methyltransferase subunit G (EC 2.1.1.86)

DE (N5-methyltetrahydromethanopterin--coenzyme M methyltransferase

DE subunit G).

DE MTRG.

OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).

OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;

OC Methanothermobacter.

OX NCBI\_TaxID=79929;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.

RX MEDLINE=9525265; PubMed=7737157;

RA Harms U., Weiss D.S., Gaertner P., Linder D., Thauer R.K.;

RT "The energy conserving N5-methyltetrahydromethanopterin:coenzyme M

methyltransferase complex from Methanobacterium thermoautotrophicum

is composed of eight different subunits.";

RL Eur. J. Biochem. 228:640-648(1995).

CC -!- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN

METHANOGENESIS. THE FORMATION OF METHYL-COENZYME M AND

TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-

TETRAHYDROMETHANOPTERIN. THIS IS A ENERGY-CONSERVING, SODIUM-ION

TRANSLLOCATING STEP.

CC -!- CATALYTIC ACTIVITY: 5-methyl-5,6,7,8-tetrahydromethanopterin + 2-

mercaptoethanesulfonate -> 5,6,7,8-tetrahydromethanopterin + 2-

(methylthio)ethanesulfonate.

CC -!- SUBUNIT: COMPOSED OF 8 DIFFERENT SUBUNITS.

CC -!- SURCELLULAR LOCATION: Integral membrane protein (Potential).

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; X84219; CAA59002.1; -

KW Transferase; Methyltransferase; Transmembrane; Methanogenesis.

FT INIT\_MET 0 0

FT TRANSMEM 51 71 POTENTIAL.

SQ SEQUENCE 85 AA; 9379 MW; D79D7C2BAF9619C1 CRC64;

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x MTRG\_METTM ..

Align seg 1/1 to: MTRG\_METTM from: 1 to: 85

3078 GGACTTCTGAAGGCTTCTCAAG 3055

|||||

73 GlyLeuLeuLysGlyLeuLeuLys 80

seq\_name: SwissProt\_40:DTD\_CLOBAB

seq\_documentation\_block:

ID

AC

DT

DT

DT

DE

DE

GN

OS

OC

OC

OX

RN

RP

RX

RA

RA

RA

RA

RA

RT

RT

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

DTD\_CLOBAB STANDARD; PRT; 149 AA.

Q97GU2;

01-MAR-2002 (Rel. 41, Created)

01-MAR-2002 (Rel. 41, Last sequence update)

01-MAR-2002 (Rel. 41, Last annotation update)

D-TYROSYL-tRNA(Tyr) deacylase (EC 3.1.1.-).

DTD OR CAC2273.

Clostridium acetobutylicum.

Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

Clostridium.

NCBI\_TaxID=1488;

[1]

SEQUENCE FROM N.A.

STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

MEDLINE=21359325; PubMed=11466286;

Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,

Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

Bennett G.N., Koonin E.V., Smith D.R.;

"Genome sequence and comparative analysis of the solvent-producing

bacterium Clostridium acetobutylicum.";

J. Bacteriol. 183:4823-4838(2001).

-!- FUNCTION: Hydrolyzes D-tyrosyl-tRNA(Tyr) into D-tyrosine and free

tRNA(Tyr). Could be a defense mechanism against a harmful effect

of D-tyrosine (By similarity).

-!- SURCELLULAR LOCATION: Cytoplasmic (Probable).

-!- SIMILARITY: BELONGS TO THE DTD FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AE007728; AAK80230.1; -

InterPro: IPR003732; DUF154.

Pfam: PF02580; DUF154; 1.

Hydrolase; Complete proteome.

SEQUENCE 149 AA; 16586 MW; 69B03FDDC1A8791E CRC64;

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x DTD\_CLOBAB ..

Align seg 1/1 to: DTD\_CLOBAB from: 1 to: 149

773 AGTCTGTTAAAGTAGATGGCAA 750

|||||

11 SerSerValLysValAspGlyLys 18

seq\_name: SwissProt\_40:Y168\_ADE02

seq\_documentation\_block:

Y168\_ADE02 STANDARD; PRT; 168 AA.

AC P03292;

21-JUL-1986 (Rel. 01, Created)

21-JUL-1986 (Rel. 01, Last sequence update)

01-APR-1990 (Rel. 14, Last annotation update)

Hypothetical protein C-168.

Human adenovirus type 2.

Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

NCBI\_TaxID=10515;

[1]

SEQUENCE FROM N.A.

MEDLINE=83056843; PubMed=7142161;

Gingeras T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,



```
RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
RT "Nucleotide sequences from the adenovirus-2 genome.";
RL J. Biol. Chem. 257:13475-13491(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83056844; PubMed=7142162;
RA Alestroem P., Akusjærvi G., Pettersson M., Pettersson U.;
RT "DNA sequence analysis of the region encoding the terminal protein
and the hypothetical N-gene product of adenovirus type 2.";
RL J. Biol. Chem. 257:13492-13498(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01917; ; NOT_ANNOTATED_CDS.
DR PIR; A03864; A03864.
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 17703 MW; F29A2BDCEE95F80E CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Y168_ADE02 ..
Align seg 1/1 to: Y168_ADE02 from: 1 to: 168

3074 AGTCGAAGTCTCCAGCTCCGGA 3097
|||||
29 SerProSerSerSerSerGly 36

seq_name: SwissProt_40:Y105_MYCPN

seq_documentation_block:
ID Y105_MYCPN STANDARD; PRT; 202 AA.
AC P75528;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG105 homolog (K04_orf202).
GN MPN244 OR MP588.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000057; AAB96236.1; -
DR InterPro; IPR003390; DUF147.
```

```
DR Pfam; PF02457; DUF147; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 6 26 POTENTIAL.
SQ SEQUENCE 202 AA; 22733 MW; F2E3F12431ADCE18 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Y105_MYCPN ..
Align seg 1/1 to: Y105_MYCPN from: 1 to: 202

1286 TTCCTGATCTTGGCAGCTTAACCTTA 1309
|||||
16 PheLeuIleLeuAlaLeuThrLeu 23

seq_name: SwissProt_40:RS4E_THEAC

seq_documentation_block:
ID RS4E_THEAC STANDARD; PRT; 235 AA.
AC Q56230;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S4e.
GN RPS4E OR TAI259.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
OC Thermoplasma
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RA Thomas N.A., Jarrell K.F.;
RT "Nucleotide sequence of the ribosomal protein genes S4e and L5 in the
archaeon Thermoplasma acidophilum.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
[2]

seq_name: SwissProt_40:Y105_MYCPN

seq_documentation_block:
ID Y105_MYCPN STANDARD; PRT; 202 AA.
AC P75528;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG105 homolog (K04_orf202).
GN MPN244 OR MP588.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57643; AAB02244.1; ALT_INIT.
DR EMBL; AL445067; CAC12383.1; ALT_INIT.
DR InterPro; IPR000876; Ribosomal_S4e.
DR InterPro; IPR002942; S4.
DR Pfam; PF00900; Ribosomal_S4e; 1.
DR Pfam; PF01479; S4; 1.
DR ProDom; PD002667; Ribosomal_S4e; 1.
DR SMART; SM00363; S4; 1.
DR PROSITE; PS00528; RIBOSOMAL_S4e; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 235 AA; 26265 MW; E3B81B6B1CA123C6 CRC64;

alignment_scores:
```

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2/rev x RS4E\_THEAC ..

Align seg 1/1 to: RS4E\_THEAC from: 1 to: 235

767 GTTAAAGTAGTGCAAAACAGTA 744

|||||

66 VallysValaspGlylysthrVal 73

seq\_name: SwissProt\_40:Y24K\_STRGR

## seq\_documentation\_block:

ID Y24K\_STRGR STANDARD; PRT; 238 AA.

AC P12752;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 24.7 kDa protein in photolysate 5' region.

OS Streptomyces griseus.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=1911;

RN [1]

.RP SEQUENCE FROM N.A.

RA MEDLINE=89315214; PubMed=2501760;

RA Kobayashi T., Takao M., Okawa A., Yasui A.;

RT "Molecular characterization of a gene encoding a photolysate from

Streptomyces griseus";

RL Nucleic Acids Res. 17:4731-4744(1989).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X15060; CAA33160.1; .

DR PIR; S05572; S05572.

KW Hypothetical protein.

SQ SEQUENCE 238 AA; 24655 MW; 046D5152D859A4ED CRC64;

## alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2/rev x Y24K\_STRGR ..

Align seg 1/1 to: Y24K\_STRGR from: 1 to: 238

1932 CATGTCCTGCTGCGTGAGAGA 1909

|||||

72 HisGlyproAlaGlyGlyGlyArg 79

seq\_name: SwissProt\_40:VGLL\_MCMVK

## seq\_documentation\_block:

ID VGLL\_MCMVK STANDARD; PRT; 274 AA.

AC P52513;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Glycoprotein L precursor.

GN GL OR UL115.

OS Murine cytomegalovirus (strain K181).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Muromegalovirus.  
OX NCBI\_TaxID=69156;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95053910; PubMed=7964634;  
RA Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,  
RA Shellam G.R.;  
RT "Identification, sequencing and expression of the glycoprotein L gene  
RT of murine cytomegalovirus";  
RL J. Gen. Virol. 75:3235-3240(1994).  
CC -!- FUNCTION: ASSOCIATED WITH GLYCOPROTEIN H (GH) TO FORM A COMPLEX  
CC IMPORTANT FOR INFECTION AND CELL FUSION. THIS ASSOCIATION IS  
CC NECESSARY FOR THE CORRECT PROCESSING AND CELL SURFACE EXPRESSION  
CC OF GH (BY SIMILARITY).  
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN L.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L32962; AAA57343.1; .

DR InterPro; IPR002689; Cytomegalo\_gL.

DR Pfam; PF01801; Cytomegalo\_gL; 1.

KW Glycoprotein; Signal; Envelope protein.

FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 18 274 GLYCOPROTEIN L.

FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 274 AA; 31239 MW; 504D349FDBEC5C03 CRC64;

## alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x VGLL\_MCMVK ..

Align seg 1/1 to: VGLL\_MCMVK from: 1 to: 274

1684 AATGCGAGTGAATTCACCTTTA 1707

|||||

64 AsnGlySerGluPheGlnProLeu 71

seq\_name: SwissProt\_40:VGLL\_MCMVS

## seq\_documentation\_block:

ID VGLL\_MCMVS STANDARD; PRT; 274 AA.

AC P52514;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Glycoprotein L precursor.

GN GL OR UL115.

OS Murine cytomegalovirus (strain Smith).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Muromegalovirus.

OX NCBI\_TaxID=10367;

RN [1]

.RP SEQUENCE FROM N.A.

RX MEDLINE=95053910; PubMed=7964634;

RA Xu J., Scalzo A.A., Lyons P.A., Farrell H.E., Rawlinson W.D.,

RA Shellam G.R.;

RT "Identification, sequencing and expression of the glycoprotein L gene

```
RT of murine cytomegalovirus."
RL J. Gen. Virol. 75:3235-3240(1994).
CC -!- FUNCTION: ASSOCIATED WITH GLYCOPROTEIN H (GH) TO FORM A COMPLEX
CC IMPORTANT FOR INFECTION AND CELL FUSION. THIS ASSOCIATION IS
CC NECESSARY FOR THE CORRECT PROCESSING AND CELL SURFACE EXPRESSION
CC OF GH (BY SIMILARITY)
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN L.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32963; AAA57344.1; -
DR InterPro; IPR002689; Cytomegalo_gL.
DR Pfam; PF01801; Cytomegalo_gL; 1.
KW Glycoprotein; Signal; Envelope protein.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 274 GLYCOPROTEIN L.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 274 AA; 31221 MW; 504D359EDBED5D03 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x VGLL_MCMVS ..
Align seg 1/1 to: VGLL_MCMVS from: 1 to: 274

1684 AATGGCAGTGAATCCCAACTTTA 1707
|||||
64 AsnGlySerGluPheGlnProLeu 71

seq_name: SwissProt_40:YOR5_ADEG1

seq_documentation_block:
ID YOR5_ADEG1 STANDARD; PRT; 283 AA.
AC P20747;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 31.5 kDa protein (ORF 5) (ORF10).
OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX NCBI_TaxID=10553;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90251474; PubMed=2160072;
RA Akopian T.A., Kruglyak V.A., Rivkina M.B., Naroditsky B.S.,
RA Tikhonenko T.I.;
RA "Sequence of an avian adenovirus (CELO) DNA fragment (0-11.2%).";
RL Nucleic Acids Res. 18:2825-2825(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96186720; PubMed=8627769;
RA Chioocca S., Kurzbaue R., Schaffner G., Baker A., Mautner V.,
RA Cotten M.;
RA "The complete DNA sequence and genomic organization of the avian
RT adenovirus CELO."
RL J. Virol. 70:2939-2949(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17217; CAA35087.1; -
DR EMBL; U46933; AAC54931.1; -
DR PIR; S10005; S10005.
KW Hypothetical protein.
SQ SEQUENCE 283 AA; 31487 MW; 6018412DA598183D CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x YOR5_ADEG1 ..
Align seg 1/1 to: YOR5_ADEG1 from: 1 to: 283

286 ACGAACCTCACCTCACCATTAAAC 309
|||||
222 ThrAsnLeuThrLeuThrIleasn 229

seq_name: SwissProt_40:PDXK_SALTY

seq_documentation_block:
ID PDXK_SALTY STANDARD; PRT; 288 AA.
AC P40192;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Pyridoxine kinase (EC 2.7.1.35) (pyridoxal kinase) (Vitamin B6
DE kinase) (Pyridoxamine kinase) (PN/PL/PM kinase).
GN PDXK OR STM2435.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=95337418; PubMed=7612925;
RA Titgemeyer F.M., Reizer J., Reizer A., Tang J., Parr T.R. Jr.,
RA Saier M.H. Jr.;
RT "Nucleotide sequence of the region between crr and cysM in Salmonella
RT typhimurium: five novel ORFs including one encoding a putative
RT transcriptional regulator of the phosphotransferase system.";
RL DNA Seq. 5:145-152(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: PHOSPHORYLATE B6 VITAMERS; FUNCTIONS IN A SALVAGE
CC PATHWAY. USES PYRIDOXAL, PYRIDOXINE, AND PYRIDOXAMINE AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: ATP + pyridoxal = ADP + pyridoxal 5'-
CC phosphate.
CC -!- COFACTOR: ZINC OR MAGNESIUM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PYRIDOXINE KINASE FAMILY.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; U11243; AAC43343.1; -.  
DR EMBL; AE008809; AAL21329.1; -.  
DR StyGene; SGI0475; pdxk.  
DR InterPro; IPR002173; PfkB.  
DR Pfam; PF00294; pfkB; 1.  
KW Transferase; Kinase; Zinc; Magnesium; Complete proteome.  
FT CONFLICT 241 241 A -> P (IN REF. 1).  
FT CONFLICT 283 283 MISSING (IN REF. 1).  
SQ SEQUENCE 288 AA; 30967 MW; CF3F701FEA841F7A CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x PDXK\_SALTY ..

Align seg 1/1 to: PDXK\_SALTY from: 1 to: 288

730 GTCACAGCCGCTCCCTACTGTTTG 753

|||||

51 ValThrAlaValProThrValLeu 58

seq\_name: SwissProt\_40:VIME\_BOVIN

seq\_documentation\_block:

ID VIME\_BOVIN STANDARD; PRT; 465 AA.

AC P48616;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Vimentin.

OS VIM.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OC NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94193008; PubMed=8144034;

RA Hess J.F., Casseleman J.T., FitzGerald P.G.;

RT "Nucleotide sequence of the bovine vimentin-encoding cDNA.";

RL Gene 140:257-259(1994).

CC -!- FUNCTION: VIMENTIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN

CC VARIOUS NONEPITHELIAL CELLS, ESPECIALLY MESENCHYMAL CELLS.

CC -!- SUBUNIT: HOMOPOLYMER.

CC -!- PTM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF

CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL

CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY

CC REORGANIZED (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; LI3263; AAA53661.1; -.

DR HSSP; P03069; 1SWI.

DR InterPro; IPR001664; IF.  
DR Pfam; PF00038; filament; 1.  
DR PROSITE; PS00226; IF; 1.  
KW Intermediate filament; Coiled coil; Phosphorylation.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT DOMAIN 1 94 HEAD.  
FT DOMAIN 95 406 ROD.  
FT DOMAIN 407 465 TAIL.  
FT DOMAIN 95 130 COIL 1A.  
FT DOMAIN 131 152 LINKER 1.  
FT DOMAIN 153 244 COIL 1B.  
FT DOMAIN 245 267 LINKER 12.  
FT DOMAIN 268 406 COIL 2.  
FT MOD\_RES 38 38 PHOSPHORYLATION (BY CAM-KINASE II)  
FT MOD\_RES 82 82 (BY SIMILARITY).  
FT MOD\_RES 82 82 PHOSPHORYLATION (BY CAM-KINASE II)  
FT MOD\_RES 82 82 (BY SIMILARITY).  
SQ SEQUENCE 465 AA; 53545 MW; 119E126778BF5801 CRC64;

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x VIME\_BOVIN ..

Align seg 1/1 to: VIME\_BOVIN from: 1 to: 465

542 TCAGCCTTGAGGCCAACACATCT 565

|||||

41 SerAlaLeuArgProThrThrSer 48

seq\_name: SwissProt\_40:GLGA\_HAEIN

seq\_documentation\_block:

ID GLGA\_HAEIN STANDARD; PRT; 476 AA.

AC P45179;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Glycogen synthase (EC 2.4.1.21) (Starch [bacterial glycogen]

DE synthase).

GN GLGA OR H11360.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OC NCBI\_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RD / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fieischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Kienle A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus

RT influenzae Rd.";

RL Science 269:496-512(1995).

CC -!- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.

CC -!- CATALYTIC ACTIVITY: ADP-glucose + [(1,4)-alpha-D-glucosyl](N) -

CC ADP + [(1,4)-alpha-D-glucosyl](N+1).

CC -!- PATHWAY: Glycogen biosynthesis; second step.

CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE

CC FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; U32815; AAC23007.1; -  
DR TIGR; H11360; -  
DR InterPro: IPR001296; Glycos\_transf\_1.  
DR Pfam: PF00534; Glycos\_transf\_1; 1.  
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;  
KW Complete proteome.  
FT BINDING 15 15 ADP-GLUCOSE (BY SIMILARITY).  
SQ SEQUENCE 476 AA; 52975 MW; E4060E820E5D768D CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x GLGA\_HAEIN ..

Align seg 1/1 to: GLGA\_HAEIN from: 1 to: 476

1524 TTGTGTTTGAATCTCAACTCCT 1501  
|||||  
421 PheValPheIleuSerAlaThrPro 428

seq\_name: SwissProt\_40:COBQ\_METTH

seq\_documentation\_block:

ID COBQ\_METTH STANDARD; PRT: 504 AA.

AC O26880;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable cobryic acid synthase.  
GN COBQ OR MTH787.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanothermobacter.  
OX NCBI\_TaxID=145262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DELTA H;

RX MEDLIN=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT deltaH: functional analysis and comparative genomics."  
RL J. Bacteriol. 179:7135-7155(1997).

CC -!- FUNCTION: CATALYZES AMIDATIONS AT POSITIONS B, D, E, AND G ON  
CC ADENOSYLCOBYRINIC A.C-DIAMIDE. NH(2) GROUPS ARE PROVIDED BY  
CC GLUTAMINE, AND ONE MOLECULE OF ATP IS HYDROGENOLYZED FOR EACH  
CC AMIDATION (BY SIMILARITY).  
CC -!- PATHWAY: COBALAMIN BIOSYNTHESIS.

CC -!- SIMILARITY: BELONGS TO THE COBB/COBQ FAMILY. COBQ SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE000857; AAB85289.1; ALT\_INIT.  
DR InterPro: IPR002586; CbIA.  
DR Pfam: PF01656; CbIA; 1.  
KW Cobalamin biosynthesis; Porphyrin biosynthesis; Complete proteome.  
SQ SEQUENCE 504 AA; 55831 MW; 8D4ABA35487A501E CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x COBQ\_METTH ..

Align seg 1/1 to: COBQ\_METTH from: 1 to: 504

194 TCACAGGCAGAGTTTGTAGGAAC 171  
|||||  
377 SerGlnGlyArgValLeuGlyAsn 384

seq\_name: SwissProt\_40:KPYG\_TOBAC

seq\_documentation\_block:

ID KPYG\_TOBAC STANDARD; PRT: 562 AA.

AC Q40546;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Pyruvate kinase isozyme G, chloroplast precursor (EC 2.7.1.40).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. PETIT HAVANA SRI; TISSUE=Seed;  
RX MEDLINE=95170010; PubMed=7865798;  
RA Blakeley S.D., Gottlob-McHugh S., Wan J., Crews L., Miki B., Ko K.,  
RA Dennis D.T.;  
RT "Molecular characterization of plastid pyruvate kinase from castor  
RT and tobacco".

RL Plant Mol. Biol. 27:79-89(1995).  
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate = ADP + phosphoenolpyruvate.  
CC -!- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM (BY SIMILARITY).

CC -!- PATHWAY: FINAL STEP IN GLYCOLYSIS.

CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Chloroplast.

CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN LEAVES. ALSO FOUND IN

CC STEMS, ROOTS AND FLOWERS.

CC -!- DEVELOPMENTAL STAGE: MOST ABUNDANTLY EXPRESSED DURING THE EARLY

CC GLOBULAR TO EARLY COTYLEDONARY STAGES OF EMBRYO DEVELOPMENT.

CC -!- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z28374; CAA82223.1; -

DR HSSP; P14178; IPKY.

DR InterPro: IPR001697; Pyruvate\_kinase.

DR Pfam: PF00224; PK; 1.

DR Pfam: PF02887; PK-C; 1.

DR PRINTS: PR01050; PYRUVTKINASE.

DR PRODOM: PD001009; Pyruvate\_kinase; 1.

DR PROSITE: PS00110; PYRUVATE\_KINASE; 1.

DR Transferase; Pyruvate; Kinase; Glycolysis; Magnesium; Chloroplast;  
KW

KW Transit peptide.  
FT TRANSIT 1 ?  
FT CHAIN 562 CHLOROPLAST (POTENTIAL).  
FT ACT\_SITE 306 PYRUVATE KINASE ISOZYME G.  
FT METAL 308 BY SIMILARITY.  
FT METAL 308 MAGNESIUM (POTENTIAL).  
FT METAL 329 MAGNESIUM (POTENTIAL).  
FT METAL 330 MAGNESIUM (POTENTIAL).  
SQ SEQUENCE 562 AA; 61867 MW; D21D50204AAG65E CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x KPYG\_TOBAC ..

Align seg 1/1 to: KPYG\_TOBAC from: 1 to: 562

627 AAACGTGTTATTATCGAAATCTTG 650  
|||||  
522 LysLeuLeuSerLysSerLeu 529

seq\_name: SwissProt\_40:PAL3\_PHAVU

seq\_documentation\_block:

ID PAL3\_PHAVU STANDARD; PRT; 710 AA.

AC P19143;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Phenylalanine ammonia-lyase class III (EC 4.3.1.5).

OS Phaseolus vulgaris (Kidney bean) (French bean)

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucotids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.

OX NCBI\_TaxID=3885;

RN [1]

RP SEQUENCE FROM N.A.

RA Cramer C.L., Edwards K., Dron M., Liang X., Dildine S.L.,

RA Bolwell G.P., Dixon R.A., Lamb C.J., Schuch W.

RT "Phenylalanine ammonia-lyase gene organisation and structure."

RL Plant Mol. Biol. 12:367-383(1989).

CC -!- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE

CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE

CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.

CC -!- CATALYTIC ACTIVITY: L-phenylalanine -> trans-cinnamate + NH(3).

CC -!- PATHWAY: KEY ENZYME OF PHENYLPROPANOID METABOLISM.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -!- PTM: CONTAINS AN ACTIVE SITE 4-METHYLIDENE-IMIDAZOLE-5-ONE (MIO),

CC WHICH IS FORMED AUTOCATALYTICALLY BY CYCLIZATION AND DEHYDRATION

CC OF RESIDUES ALA-SER-GLY (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.

CC PIR; S04128; S04128.

DR HSSP; P21310; 1B8F.

DR InterPro; IPR001106; PAL.

DR Pfam; PF00221; PAL; 1.

DR PROSITE; PS00488; PAL\_HISTIDASE; 1.

KW Lysase; Phenylpropanoid metabolism; Multigene family.

FT SITE 198 MODIFIED TO FORM 4-METHYLIDENE-IMIDAZOLE-

FT SITE 200 5-ONE (BY SIMILARITY).

SQ SEQUENCE 710 AA; 77939 MW; C170771F4AE54773 CRC64;

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x PAL3\_PHAVU ..

Align seg 1/1 to: PAL3\_PHAVU from: 1 to: 710

2144 TCAATCTTTTCACTTGAAGAAC 2167  
|||||  
436 SerAsnLeuSerValGlyArgAsn 443

seq\_name: SwissProt\_40:TLR6\_HUMAN

seq\_documentation\_block:

ID TLR6\_HUMAN STANDARD; PRT; 796 AA.

AC Q9Y2C9;

DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Toll-like receptor 6 precursor.

GN TLR6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=99250250; PubMed=10231569;

RA Takeuchi O., Kawai T., Sanjo H., Copeland N.G., Gilbert D.J.,

RA Jenkins N.A., Takeda K., Akira S.

RT "TLR6: A novel member of an expanding Toll-like receptor family."

RL Gene 231:59-65(1999).

RN [2]

RP FUNCTION.

RX PubMed=11441107;

RA Bulut Y., Faure E., Thomas L., Equils O., Arditi M.

RT "Cooperation of Toll-like receptor 2 and 6 for cellular activation by

RT soluble tuberculosis factor and Borrelia burgdorferi outer surface

RT protein A lipoprotein: role of toll-interacting protein and IL-1

RT receptor signaling molecules in Toll-like receptor 2 signaling."

RL J. Immunol. 167:987-994(2001).

CC -!- FUNCTION: Participates in the innate immune response to Gram-

CC positive bacteria and fungi. Acts via MyD88 and TRAF6, leading to

CC NF-kappa-B activation, cytokine secretion and the inflammatory

CC response. Recognizes mycoplasma macrophage-activating

CC lipopeptide-2kD (MALP-2), soluble tuberculosis factor (STF),

CC phenol-soluble modulin (PSM) and B.burgdorferi outer surface

CC protein A lipoprotein (OsPA-L) cooperatively with TLR2.

CC -!- SUBUNIT: Binds TLR2 via their respective extracellular domains.

CC Binds MyD88 via their respective TIR domains (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Plasma membrane and

CC phagosomes (By similarity).

CC -!- TISSUE SPECIFICITY: Detected in monocytes, CD11c+ immature

CC dendritic cells, plasmacytoid pre-dendritic cells and dermal

CC microvessel endothelial cells.

CC -!- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.

CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.

CC -!- SIMILARITY: CONTAINS 13 LEUCINE-RICH REPEATS (LRR).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AB020807; BAA78631.1; ..

CC MIM; 605403; ..

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR000157; TIR.

DR Pfam; PF00560; LRR; 8.

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01582; TIR; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 2.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00255; TIR; 1.  
DR PROSITE; PS0104; TIR; 1.  
KW Receptor; Immune response; Inflammatory response; Signal;  
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.  
FT SIGNAL 1 31  
FT CHAIN 32 796  
FT DOMAIN 32 796  
FT TRANSMEM 587 607  
FT DOMAIN 608 796  
FT REPEAT 51 74  
FT REPEAT 75 98  
FT REPEAT 100 120  
FT REPEAT 121 144  
FT REPEAT 158 175  
FT REPEAT 176 201  
FT REPEAT 222 248  
FT REPEAT 376 400  
FT REPEAT 402 428  
FT REPEAT 430 447  
FT REPEAT 449 472  
FT REPEAT 474 496  
FT REPEAT 498 517  
FT REPEAT 517 784  
FT DOMAIN 640 784  
FT CARBOHYD 144 144  
FT CARBOHYD 186 186  
FT CARBOHYD 214 214  
FT CARBOHYD 253 253  
FT CARBOHYD 285 285  
FT CARBOHYD 359 359  
FT CARBOHYD 423 423  
FT CARBOHYD 434 434  
FT CARBOHYD 563 583  
FT CARBOHYD 583 583  
SQ SEQUENCE 796 AA; 91889 MW; 35CEAEC05BFA8BD CRC64;  
  
alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x TLR6\_HUMAN ..  
Align seg 1/1 to: TLR6\_HUMAN from: 1 to: 796  
2194 TTGGACCTCAGCCACACCAACCTG 2217  
|||||  
105 LeuAspLeuSerHisAsnGlnLeu 112  
  
seq\_name: SwissProt\_40:BIMA\_EMENI  
  
seq\_documentation\_block:  
ID BIMA\_EMENI STANDARD; PRT; 806 AA..  
AC P17885;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein bima.  
GN BIMA.  
OS Emericella nidulans (Aspergillus nidulans).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiales; Trichocomaceae; Emericella.  
OX NCBI\_TaxID=5072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R153;  
RA O'Donnell K.L.; Osmani A.H.; Osmani S.A.; Morris N.R.;  
RT "bima encodes a member of the tetratricopeptide repeat family of  
RT proteins and is required for the completion of mitosis in Aspergillus  
RT nidulans.";

RL J. Cell Sci. 99:711-719(1991).  
CC -!- FUNCTION: REQUIRED FOR THE COMPLETION OF MITOSIS IN ASPERGILLUS  
CC NIDULANS.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- SIMILARITY: BELONGS TO THE CDC27/NUC2 FAMILY.  
CC -!- SIMILARITY: CONTAINS 8 TPR REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X59269; CAA41959.1; ..  
DR PIR; S21860; S21860.  
DR PIR; A53256; A53256.  
DR InterPro; IPR001440; TPR.  
DR Pfam; PF00515; TPR; 7.  
DR SMART; SM00028; TPR; 7.  
KW Cell division; Cell cycle; Mitosis; Repeat; TPR repeat;  
KW Nuclear protein.  
FT REPEAT 76 109 TPR 1.  
FT REPEAT 127 160 TPR 2.  
FT DOMAIN 260 399 BIMA DOMAIN.  
FT REPEAT 513 546 TPR 3.  
FT REPEAT 581 614 TPR 4.  
FT REPEAT 616 648 TPR 5.  
FT REPEAT 649 682 TPR 6.  
FT REPEAT 684 716 TPR 7.  
FT REPEAT 751 784 TPR 8.  
SQ SEQUENCE 806 AA; 89714 MW; F137BDE3A74C0457 CRC64;  
  
alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x BIMA\_EMENI ..  
Align seg 1/1 to: BIMA\_EMENI from: 1 to: 806  
946 GCGCTGACAGCAATTAAGTTTAA 969  
|||||  
736 AlaLeuThrGluLeuLysValLeu 743  
  
seq\_name: SwissProt\_40:CYP1\_BRUMA  
  
seq\_documentation\_block:  
ID CYP1\_BRUMA STANDARD; PRT; 843 AA.  
AC Q27450;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Peptidylprolyl isomerase CYP-1 (EC 5.2.1.8) (Peptidylprolyl cis-trans  
DE isomerase) (Cyclophilin) (PPIase).  
GN CYP.  
OS Brugia malayi.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
OC Onchocercidae; Brugia.  
OX NCBI\_TaxID=6279;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96009642; PubMed=7547885;  
RA Page A.P.; Landry D.; Wilson G.G.; Carlow C.K.S.;  
RT "Molecular characterization of a cyclosporin A-insensitive  
RT cyclophilin from the parasitic nematode Brugia malayi.";  
RT Biochemistry 34:11545-11550(1995).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF 1-177.

RX MEDLINE-98218582; PubMed-9559680;  
RA Taylor P., Page A.P., Kontopidis G., Husi H., Walkinshaw M.D.;  
RT "The X-ray structure of a divergent cyclophilin from the nematode  
parasite Brugia malayi.";  
RL FEBS Lett. 425:361-366(1998).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 1-177.  
RC STRAIN-ATCC 75593;  
RX MEDLINE-98318040; PubMed-9655334;  
RA Mikol V., Ma D., Carlow C.K.S.;  
RT "Crystal structure of the cyclophilin-like domain from the parasitic  
nematode Brugia malayi.";  
RL Protein Sci. 7:1310-1316(1998).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.47 ANGSTROMS) OF 1-177.  
RX MEDLINE-20108543; PubMed-10642184;  
RA Ellis P.J., Carlow C.K.S., Ma D., Kuhn P.;  
RT "Crystal structure of the complex of Brugia malayi cyclophilin and  
cyclosporin A.";  
RL Biochemistry 39:592-598(2000).  
CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC  
PEPTIDE BONDS IN OLIGOPEPTIDES.  
CC -1- ENZYME REGULATION: RELATIVELY INSENSITIVE TO INHIBITION BY CSA.  
CC -1- SIMILARITY: BELONGS TO THE CYCLOPHILIN-TYPE PPIASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; I37292; AAC37249.1; -  
DR PDB; 1A58; 27-MAY-98.  
DR PDB; 1A33; 29-JUL-98.  
DR PDB; 1C5F; 02-DEC-99.  
DR InterPro; IPR002130; CSA\_PPIase.  
DR Pfam; PF00160; pro\_isomerase; 1.  
DR PRINTS; PR00153; CSAPPISMASE.  
DR PROSITE; PS00170; CSA\_PPIASE\_1; 1.  
DR PROSITE; PS50072; CSA\_PPIASE\_2; 1.  
KW Isomerase; Rotamase; 3D-structure.  
FT DOMAIN 10 175 PPIASE, CYCLOPHILIN-TYPE.  
FT DOMAIN 700 709 POLY-ARG.  
FT DOMAIN 713 716 POLY-ARG.  
FT DOMAIN 800 815 POLY-SER.  
FT DOMAIN 828 837 POLY-SER.  
SQ SEQUENCE 843 AA; 97817 MW; 3C34EC90A32EDBDC CRC64;  
  
alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x CYP1\_BRUMA ..  
  
Align seg 1/1 to: CYP1\_BRUMA from: 1 to: 843  
  
3071 AGAAGTCCAGTTCCTCCAGCTCC 3094  
|||||  
799 ArgSerProSerSerSerSerSer 806  
  
seq\_name: SwissProt\_40:POL\_BLVAV  
  
seq\_documentation\_block:  
ID POL\_BLVAV STANDARD; PRT; 852 AA.  
AC P25059;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
  
seq\_name: SwissProt\_40:POL\_BLVAV  
  
seq\_documentation\_block:  
ID POL\_BLVAV STANDARD; PRT; 852 AA.  
AC P25059;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE POL polyprotein [Contains: Reverse transcriptase (EC 2.7.7.49);  
DE Endonuclease].  
GN POL.  
OS Bovine leukemia virus (Australian isolate) (BLV).  
OC Viruses; Retrovirdae; Retroviridae; Deltaretrovirus.  
OX NCBI\_TaxID=11903;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-90362060; PubMed-2167927;  
RA Coulston J., Naif H., Brandon R., Kumar S., Khan S., Daniel R.C.W.,  
RA Lavin M.F.;  
RT "Molecular cloning and sequencing of an Australian isolate of  
proliferative bovine leukaemia virus DNA: comparison with other  
isolates.";  
RL J. Gen. Virol. 71:1737-1746(1990).  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: VERY STRONG, WITH BOVINE LEUKEMIA VIRUS ENV  
POLYPROTEIN FROM OTHER ISOLATES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D00647; BAA00544.1; -  
DR FIR; J00555; GNLJGA.  
DR HSP; P03366; IHMV.  
DR InterPro; IPR001037; Integrase\_C.  
DR InterPro; IPR002156; RNaseH.  
DR InterPro; IPR000477; RVTse.  
DR InterPro; IPR001584; Rve.  
DR Pfam; PF00552; Integrase; 1.  
DR Pfam; PF00075; rnaseh; 1.  
DR Pfam; PF00665; rve; 1.  
DR Pfam; PF00078; rvt; 1.  
KW Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;  
KW Endonuclease; Polyprotein.  
SQ SEQUENCE 852 AA; 95441 MW; CD69FE59E1208A08 CRC64;  
  
alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2/rev x POL\_BLVAV ..  
  
Align seg 1/1 to: POL\_BLVAV from: 1 to: 852  
  
3151 GGTGAGCTTCGGGTTGTGGCC 3128  
|||||  
471 GlyGluLeuAlaGlyLeuLeuAla 478  
  
seq\_name: SwissProt\_40:POL\_BLVJ  
  
seq\_documentation\_block:  
ID POL\_BLVJ STANDARD; PRT; 852 AA.  
AC P03361;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE POL polyprotein [Contains: Reverse transcriptase (EC 2.7.7.49);  
DE Endonuclease].  
GN POL.  
OS Bovine leukemia virus (Japanese isolate BLV-1) (BLV).  
OC Viruses; Retrovirdae; Retroviridae; Deltaretrovirus.  
OX NCBI\_TaxID=11907;  
RN [1]  
RP SEQUENCE FROM N.A.



```
RX MEDLINE-95140159; PubMed=2983308;
RA Sagata N., Yasunaga T., Tsuzuku-Kawamura J., Ohishi K., Ogawa Y.,
RA Ikawa Y.;
RT "Complete nucleotide sequence of the genome of bovine leukemia virus:
RT its evolutionary relationship to other retroviruses.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:677-681(1985).
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: VERY STRONG, WITH BOVINE LEUKEMIA VIRUS ENV
CC POLYPROTEIN FROM OTHER ISOLATES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01210; AAA42785.1; -.
DR PIR; A03960; GNLJGB.
DR HSSP; P03366; 1HMV.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR000477; RVse.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; rvt; 1.
DR Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
KW Endonuclease; Polyprotein.
SQ SEQUENCE 852 AA; 95182 MW; A7A63524A18C458D CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x POL_BLVJ ...
Align seg 1/1 to: POL_BLVJ from: 1 to: 852
3151 GGTGAGCTTGGCGTTTGTGGCC 3128
|||||
471 GlyGluLeuAlaGlyLeuLeuAla 478

seq_name: SwissProt_40:MMLA_MYCTU
seq_documentation_block:
ID MMLA_MYCTU STANDARD; PRT; 1002 AA.
AC O50439;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative membrane protein mmpL10.
GN MmpL10 OR RV1183 OR MT1220 OR MTV005.19.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV.
RX MEDLINE-98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S.,
RA
```

```
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE MMLA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL010186; CAA15860.1; -.
DR EMBL; AE006999; AAK45477.1; -.
DR TIGR; MT1220; -.
DR TubercuList; Rv1183; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
SQ SEQUENCE 1002 AA; 106414 MW; 5B49B70AF2076AD8 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x MMLA_MYCTU ..
Align seg 1/1 to: MMLA_MYCTU from: 1 to: 1002
563 ATGTTGTTGGCTCAAGCTGAGA 540
|||||
879 MetLeuLeuAlaSerArgLeuArg 886

seq_name: SwissProt_40:MMLA_MYCLE
seq_documentation_block:
ID MMLA_MYCLE STANDARD; PRT; 1008 AA.
AC Q49619; Q9CC63;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative membrane protein mmpL10.
GN MmpL10 OR TPI OR ML1231 OR Bll170_C1_181.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
```

```

OC NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elgmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -|- SIMILARITY: BELONGS TO THE MPL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00010; AAA17062.1; -.
DR EMBL; AL583921; CAC31612.1; ALT_INIT.
DR Leproma; ML1231; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 225 245 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 340 360 POTENTIAL.
FT TRANSMEM 389 409 POTENTIAL.
FT TRANSMEM 587 607 POTENTIAL.
FT TRANSMEM 835 855 POTENTIAL.
FT TRANSMEM 862 882 POTENTIAL.
FT TRANSMEM 895 915 POTENTIAL.
FT TRANSMEM 940 960 POTENTIAL.
FT TRANSMEM 961 981 POTENTIAL.
SQ SEQUENCE 1008 AA; 106835 MW; A554907B3ADDF130 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x WMLA_MYCLE ..
Align seg 1/1 to: WMLA_MYCLE from: 1 to: 1008
563 ATGTTGTTGGCTCAAGGCTGAGA 540
|||||
912 MetLeuLeuAlaSerArgLeuArg 919
seq_name: SwissProt_40:SVI_BORBU
seq_documentation_block:
ID SVI_BORBU STANDARD; PRT; 1042 AA.
AC O51773;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT 15-DEC-1998 (Rel. 37, Created)

DE Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--tRNA ligase)
DE (IleRS).
GN ILES OR BB0833.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Wathey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.";
RL Nature 390:580-586(1997).
CC -|- CATALYTIC ACTIVITY: ATP + L-Isoleucine + tRNA(Ile) = AMP +
CC diphosphate + L-Isoleucyl-tRNA(Ile).
CC -|- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001181; AAC67179.1; -.
DR HSSP; P56690; IILE.
DR TIGR; BB0833; -.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002301; tRNA-synt_1e.
DR Pfam; PF00133; tRNA-synt_1.
DR PRINTS; PR00984; TRNASYNTHILE.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; zinc; Complete proteome.
FT SITE 48 58 "HIGH" REGION.
FT SITE 594 598 "KMSKS" REGION.
FT BINDING 597 597 ATP (BY SIMILARITY).
SQ SEQUENCE 1042 AA; 122331 MW; 6C0F7D820CA32F75 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x SVI_BORBU ..
Align seg 1/1 to: SVI_BORBU from: 1 to: 1042
2116 AACATCTCTCTTTGCCCAAAAT 2139
|||||
670 LysAsnLeuSerLeuAlaLysAsn 677
seq_name: SwissProt_40:DP3A_AQUAE
seq_documentation_block:
ID DP3A_AQUAE STANDARD; PRT; 1161 AA.
AC O67125;
DT 15-DEC-1998 (Rel. 37, Created)
```

```
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase III alpha subunit (EC 2.7.7.7).
GN DNAE OR AQ.1008.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CC CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
CC DIMERIZATION TO FORM THE POLIII' COMPLEX. POLIII' ASSOCIATES WITH
CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. DNAE
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000718; AAC07087.1; -.
DR InterPro; IPR004013; PHP_C.
DR InterPro; IPR003141; PHP_N.
DR InterPro; IPR002309; tRNA-synt.2.
DR Pfam; PF02811; PHP_C; 1.
DR Pfam; PF02231; PHP_N; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR SMART; SM00481; POLIIIac; 1.
DR Transferrase; DNA-directed DNA polymerase; DNA replication;
KW Complete proteome.
SQ SEQUENCE 1161 AA; 133207 MW; 619F7436E1262BD4 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x YAB2_SCHPO ..
Align seg 1/1 to: YAB2_SCHPO from: 1 to: 1318

3231 CTAGCCCTCTTGTGCAAAACACAA 3254
199 LeuAlaLeuLeuCysLysThrGln 206

seq_name: SwissProt_40:ICP4_HSVMG
seq_documentation_block:
ID ICP4_HSVMG STANDARD; PRT; 1415 AA.
AC Q02362;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Trans-acting transcriptional activator protein ICP4 (Immediate-early
DE protein IE175).
GN ICP4.
OS Marek's disease herpesvirus (strain GA) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10388;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92351564; PubMed=1322594;
RA Anderson A.S., Francesconi A., Morgan R.W.;
RT "Complete nucleotide sequence of the Marek's disease virus ICP4
RT gene.";
RL Virology 189:657-667(1992).
CC -1- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS. IT IS
CC REQUIRED FOR THE SWITCH FROM IMMEDIATE-EARLY TO EARLY MODE OF
CC GENE EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.
CC -1- PTM: A LONG STRETCH OF SERINE RESIDUES MAY BE A MAJOR SITE OF
CC PHOSPHORYLATION.

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 151.6 kDa protein C2G11.02 in chromosome 1.
GN SPAC2G11.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO YEAST YJRO41C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54354; CAA91167.1; -.
DR Hypothetical protein.
SQ SEQUENCE 1318 AA; 151569 MW; 4735560E45B2E03A CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x YAB2_SCHPO ..
Align seg 1/1 to: YAB2_SCHPO from: 1 to: 1318

3231 CTAGCCCTCTTGTGCAAAACACAA 3254
199 LeuAlaLeuLeuCysLysThrGln 206

seq_name: SwissProt_40:ICP4_HSVMG
seq_documentation_block:
ID ICP4_HSVMG STANDARD; PRT; 1415 AA.
AC Q02362;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Trans-acting transcriptional activator protein ICP4 (Immediate-early
DE protein IE175).
GN ICP4.
OS Marek's disease herpesvirus (strain GA) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10388;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92351564; PubMed=1322594;
RA Anderson A.S., Francesconi A., Morgan R.W.;
RT "Complete nucleotide sequence of the Marek's disease virus ICP4
RT gene.";
RL Virology 189:657-667(1992).
CC -1- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS. IT IS
CC REQUIRED FOR THE SWITCH FROM IMMEDIATE-EARLY TO EARLY MODE OF
CC GENE EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.
CC -1- PTM: A LONG STRETCH OF SERINE RESIDUES MAY BE A MAJOR SITE OF
CC PHOSPHORYLATION.

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase III alpha subunit (EC 2.7.7.7).
GN DNAE OR AQ.1008.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CC CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
CC DIMERIZATION TO FORM THE POLIII' COMPLEX. POLIII' ASSOCIATES WITH
CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. DNAE
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000718; AAC07087.1; -.
DR InterPro; IPR004013; PHP_C.
DR InterPro; IPR003141; PHP_N.
DR InterPro; IPR002309; tRNA-synt.2.
DR Pfam; PF02811; PHP_C; 1.
DR Pfam; PF02231; PHP_N; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR SMART; SM00481; POLIIIac; 1.
DR Transferrase; DNA-directed DNA polymerase; DNA replication;
KW Complete proteome.
SQ SEQUENCE 1161 AA; 133207 MW; 619F7436E1262BD4 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x DP3A_AQUAE ..
Align seg 1/1 to: DP3A_AQUAE from: 1 to: 1161

3090 CTGAGGAACTTGAGCTTCTGAA 3067
583 LeuGlulglulgluLeuLys 590

seq_name: SwissProt_40:YAB2_SCHPO
seq_documentation_block:
ID YAB2_SCHPO STANDARD; PRT; 1318 AA.
AC Q09604;
DT 01-NOV-1995 (Rel. 32, Created)
```

```
CC -|- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M75729; AAA46111.1; -.
DR PIR; A42991; EDBEGA.
KW Early protein; Transcription regulation; Trans-acting factor;
FT DNA-binding; Phosphorylation; Nuclear protein.
FT DOMAIN 155 200 SER/PRO-RICH.
SQ SEQUENCE 1415 AA; 154936 MW; C0846F7BEF4D1126 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x ICP4_HSVGM ..
Align seg 1/1 to: ICP4_HSVGM from: 1 to: 1415

3071 AGAAGTCCCAAGTCTCCAGCTCC 3094
|||||
156 ArgSerProSerSerSerSer 163

seq_name: SwissProt_40:CAT8_YEAST

seq_documentation_block:
ID CAT8_YEAST STANDARD; PRT; 1433 AA.
AC P39113;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Regulatory protein CAT8.
GN CAT8 OR MSP8 OR YMR280C OR YMR021.06C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95198709; PubMed=7891685;
RA Hedges D., Proft M., Entian K.-D.;
RT "CAT8, a new zinc cluster-encoding gene necessary for derepression of
RT glucogenic enzymes in the yeast Saccharomyces cerevisiae.";
RL Mol. Cell. Biol. 15:1915-1922(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Pearson D., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ENY.WA-1A;
RA Boles E., Hettmann C., Zimmermann F.K.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: ACTIVATOR OF THE GLUCONEOGENIC ENZYMES FBPI AND PCK1
CC GENES.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- PTM: COULD BE THE TARGET OF THE SNF1/CAT1 - SNF4/CAT3 KINASE
CC COMPLEX.
CC -|- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC CLUSTER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X78344; CAA55139.1; -.
DR EMBL; Z49704; CAA89778.1; -.
DR EMBL; X94215; CAA63906.1; -.
DR PIR; S48234; S48234.
DR HSSP; P04386; IAW6.
DR TRANSFAC; T03227; -.
DR SGD; S0004893; CAT8.
DR InterPro: IPR001138; ZN2_Cy6_FUNGAL_1;
DR Pfam; PF00172; Zn_c1us; 1.
DR PRINTS; PR00054; FUNGALZNCYS.
DR SMART; SM00066; GAL4; 1.
DR PROSITE; PS00463; ZN2_Cy6_FUNGAL_1;
DR PROSITE; PS00048; ZN2_Cy6_FUNGAL_2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Zinc;
KW Metal-binding; Phosphorylation; Carbohydrate metabolism; Activator.
FT DNA_BIND 70 97 ZN(2)-CYS(6), FUNGAL-TYPE.
FT DOMAIN 208 212 POLY-PRO.
FT DOMAIN 972 976 POLY-ASN.
FT CONFLICT 303 303 S -> A (IN REF. 3).
FT CONFLICT 747 747 K -> L (IN REF. 1).
FT CONFLICT 768 768 T -> A (IN REF. 3).
FT CONFLICT 999 1002 MISSING (IN REF. 3).
FT CONFLICT 1008 1008 G -> S (IN REF. 3).
FT CONFLICT 1016 1016 H -> Q (IN REF. 3).
FT CONFLICT 1019 1019 Q -> P (IN REF. 3).
FT CONFLICT 1061 1061 V -> M (IN REF. 3).
FT CONFLICT 1072 1072 T -> A (IN REF. 3).
FT CONFLICT 1074 1074 A -> S (IN REF. 3).
FT CONFLICT 1092 1092 N -> S (IN REF. 3).
FT CONFLICT 1100 1100 I -> V (IN REF. 3).
FT CONFLICT 1120 1120 M -> L (IN REF. 3).
FT CONFLICT 1162 1162 S -> A (IN REF. 3).
SQ SEQUENCE 1433 AA; 160485 MW; 5ED790BEF847B632 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x CAT8_YEAST ..
Align seg 1/1 to: CAT8_YEAST from: 1 to: 1433

526 CCTAGCTTACAGCTTCTCAGCCTT 549
|||||
520 ProSerLeuGlnLeuLeuSerLeu 527

seq_name: SwissProt_40:YJU7_YEAST

seq_documentation_block:
ID YJU7_YEAST STANDARD; PRT; 2014 AA.
AC P39526;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 229.9 kDa protein in NUC1-NC1 intergenic region.
GN YJL207C OR J0312 OR HRD550.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95274326; PubMed=7754713;
```

RA Purnelle B., Coster F., Goffeau A.;  
RT "The sequence of a 36 kb segment on the left arm of yeast chromosome  
RT x identifies 24 open reading frames including NUC1, PRP21 (SPP91),  
RT C0C6, CRV2, the gene for S24, a homologue to the aconitase gene ACO1  
RT and two homologues to chromosome III genes.";  
RL Yeast 10:1235-1249(1994).  
RN [2]  
RP SEQUENCE OF 1465-2014 FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=95242842; PubMed=7725802;  
RA Vandenbol M., Durand P., Bolle P.-A., Dion C., Portetelle D.,  
RA Hilger F.;  
RA "Sequence analysis of a 40.2 kb DNA fragment located near the left  
RT telomere of yeast chromosome X.";  
RL Yeast 10:1657-1662(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X77688; CAA54749.1; -;  
DR EMBL; Z34098; CAA84004.1; -;  
DR EMBL; Z49482; CAA89504.1; -;  
DR PIR; S46622; S46622.  
DR SGD; S0003743; YJL207C.  
KW Hypothetical protein.  
SQ SEQUENCE 2014 AA; 229898 MW; 6A296796F2FA5791 CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x YJ07\_YEAST ..

Align seg 1/1 to: YJ07\_YEAST from: 1 to: 2014

1805 CTTATATCCAGAACTCCAGTTG 1782

|||||

1189 Leu1leSerArgThrSerSerLeu 1196

seq\_name: SwissProt\_40:RL36\_RICPR

seq\_documentation\_block:

ID RL36\_RICPR STANDARD; PRT; 41 AA.

AC Q9ZD87;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 50S ribosomal protein L36.

GN RPMJ OR RP456.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

OX NCBI\_TaxID=782;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MADRID E;

RX MEDLINE=99039499; PubMed=9823893;

RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,

RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,

RA Eriksson A.-S., Winkler H.H., Kurland C.G.;

RA "The genome sequence of Rickettsia prowazekii and the origin of

RT mitochondria.";

RL Nature 396:133-140(1998)

CC -1- SIMILARITY: BELONGS TO THE L36P FAMILY OF RIBOSOMAL PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AJ235271; CAA14912.1; -;  
DR InterPro: IPR000473; Ribosomal\_L36.  
DR Pfam: PF00444; Ribosomal\_L36; 1.  
DR PROSITE: PS00828; RIBOSOMAL\_L36; FALSE\_NEG.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 41 AA; 4861 MW; EEB71BC7C7A5C76 CRC64;

alignment\_scores:

Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x RL36\_RICPR ..

Align seg 1/1 to: RL36\_RICPR from: 1 to: 41

1183 TCTTCACGTAAAGCCTGAAA 1203

|||||

5 SerSerLeuLySerLeuLys 11

seq\_name: SwissProt\_40:GCAD\_BACME

seq\_documentation\_block:

ID GCAD\_BACME STANDARD; PRT; 68 AA.

AC P28017;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE UDP-N-acetylglucosamine pyrophosphorylase (EC 2.7.7.23) (N-

DE acetylglucosamine-1-phosphate uridylyltransferase) (TMS protein)

DE (Fragment).

GN GCAD OR TMS.

OS Bacillus megaterium.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=1404;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PV361;

RX MEDLINE=92223102; PubMed=1373326;

RA Hudspeth D.S.S., Vary P.S.;

RT "spoVG sequence of Bacillus megaterium and Bacillus subtilis.";

RL Blochim. Biophys. Acta 1130:229-231(1992).

CC -1- CATALYTIC ACTIVITY: UTP + N-acetyl-alpha-D-glucosamine 1-phosphate

CC - diphosphate + UDP-N-acetyl-D-glucosamine.

CC -1- PATHWAY: PEPTIDOGLYCAN AND LIPOPOLYSACCHARIDE BIOSYNTHESIS.

CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NOFL FAMILY OF

CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; X62377; CAA44241.1; -;

DR PIR; S18901; S18901.

DR InterPro: IPR001451; Hexapep.transf.

DR PROSITE: PS00101; HEXAPEP\_TRANSFERASES; PARTIAL.

DR Peptidoglycan synthesis; Cell wall; Transferase;

KW Nucleotidyltransferase; Repeat.

KW

FT NON\_TER 68 68  
SQ SEQUENCE 68 AA: 7399 MW; CD40139522E5A130 CRC64;

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x GCAD\_BACME ..  
Align seg 1/1 to: GCAD\_BACME from: 1 to: 68

2587 TTAGATCTGACTAAGCTGATT 2607  
|||||  
44 LeuAspLeuThrAsnLeuIle 50

seq\_name: SwissProt\_40:UCRH\_BOVIN

seq\_documentation\_block:

ID UCRH\_BOVIN STANDARD; PRT; 78 AA.  
AC P00126;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ubiquinol-cytochrome C reductase complex 11 kDa protein (EC 1.10.2.2)  
DE (Mitochondrial hinge protein) (Cytochrome C1, nonheme 11 kDa protein)  
DE (Complex III subunit VIII).  
GN UCRH.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=83007120; PubMed=6126477;  
RA Wakabayashi S., Takeda H., Matsubara H., Kim C.H., King T.E.;  
RT "Identity of the heme-not-containing protein in bovine heart  
cytochrome c1 preparation with the protein mediating c1-c complex  
formation -- a protein with high glutamic acid content.";  
RL J. Biochem. 91:2077-2085(1982).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
RX MEDLINE=97349328; PubMed=9204897;  
RA Xia D., Yu C.A., Kim H., Xia J.Z., Kachurin A.M., Zhang L., Yu L.,  
RA Deisenhofer J.;  
RT "Crystal structure of the cytochrome bcl complex from bovine heart  
mitochondria.";  
RL Science 277:60-66(1997).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
RX MEDLINE=98316377; PubMed=9651245;  
RA Iwata S., Lee J.W., Okada K., Lee J.K., Iwata M., Rasmussen B.,  
RA Link T.A., Ramaswamy S., Jap B.K.;  
RT "Complete structure of the 11-subunit bovine mitochondrial cytochrome  
bcl complex.";  
RL Science 281:64-71(1998).  
CC -|- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C  
REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH  
IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS PROTEIN MAY  
MEDIATE FORMATION OF THE COMPLEX BETWEEN CYTOCHROMES C AND C1.  
CC -|- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2  
ferricytochrome c.  
CC -|- SUBUNIT: BCL COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS  
(CYTOCHROME B, CYTOCHROME C1, RIESKE PROTEIN), 2 CORE PROTEINS AND  
6 LOW-MOLECULAR WEIGHT PROTEINS.  
CC -|- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
CC -|- SIMILARITY: BELONGS TO THE UCRH/QCR6 FAMILY.  
DR PIR: A00119; CCB011.  
DR PDB: 1QCR; 14-OCT-98.  
DR PDB: 1BGY; 23-FEB-99.

DR PDB: 1BE3; 16-FEB-99.  
DR InterPro: IPR003422; UCR\_hinge.  
DR Pfam: PF02320; UCR\_hinge; 1.  
KW Mitochondrion; Inner membrane; Electron transport; Respiratory chain;  
KW Oxidoreductase; 3D-structure.  
FT DOMAIN 5 12 POLY-GLU.  
SQ SEQUENCE 78 AA; 9175 MW; 4A2D25A555455560 CRC64;

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2/rev x UCRH\_BOVIN ..

Align seg 1/1 to: UCRH\_BOVIN from: 1 to: 78

1247 RAGCTTTTCAACTCTTTAAAG 1227

|||||  
72 LysLeuPheAsnSerLeuLys 78







OM of: US-09-202-054-2 to: SPTRMBL19:\* out\_format : pfs

Date: Jul 17, 2002 1:39 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet\_n2p.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09202054/runat\_16072002.074914.7817/app\_query.fasta\_1.3393  
-DB=SPTRMBL19 -QFWT=fastan -SUFFIX=oli6.rspt -GAPOP=4 500  
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000  
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000  
-DELOP=6.000 -DELEX=7.000 -START=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=1000 -DOCALIGN=200 -THR\_SCORE=quality  
-THR\_MIN=6 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs -NORM-ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09202054\_ecn1\_l149 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -ECGN1\_TIMEOUT=30 -NO\_XLPXY -WAIT\_THREADS=1

Search information block:

Query: US-09-202-054-2

Query length: 3283

Database: SPTRMBL19:\*

Database sequences: 562222

Database length: 172994929

Search time (sec): 187.390000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000  
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
sp_human:Q9NYK1	+ 1049.00	1956.64	0.0	1049	Q9NYK1 homo sapiens (human).
sp_human:Q9NR98	+ 847.00	16024.94	0.0	1049	Q9NR98 homo sapiens (human).
sp_human:Q9NYK1	+ 48.00	868.87	4.5e-41	1050	Q9NYK1 mus musculus (mouse).
sp_human:Q9NYK1	+ 17.00	280.96	2.6e-08	1032	Q9NYK1 mus musculus (mouse).
sp_human:Q9H5G9	+ 12.00	193.23	0.0056	363	Q9H5G9 homo sapiens (human).
sp_human:Q9H5G9	+ 12.00	184.91	0.0048	1232	Q9H5G9 arabidopsis thaliana (n).
sp_human:Q9NYC2	+ 11.00	167.53	0.0566	975	Q9NYC2 homo sapiens (human).
sp_human:Q9NYC3	+ 11.00	167.15	0.0562	1032	Q9NYC3 homo sapiens (human).
sp_human:Q9NR96	+ 11.00	167.15	0.0562	1032	Q9NR96 homo sapiens (human).
sp_human:Q9NR97	+ 10.00	149.99	0.6617	791	Q9NR97 saccharomyces cerevisiae (n).
sp_human:Q9NR97	+ 10.00	148.12	0.6392	1041	Q9NR97 homo sapiens (human).
sp_human:Q9NYG9	+ 10.00	148.00	0.6378	1059	Q9NYG9 homo sapiens (human).
sp_human:Q9NYG9	+ 9.00	145.39	9.84	96	Q9NYG9 rattus norvegicus (rat).
sp_human:Q9P021	+ 9.00	138.87	8.72	250	Q9P021 homo sapiens (human).
sp_human:Q94169	+ 9.00	138.81	8.71	252	Q94169 potentilla anserina. pol.
sp_human:Q94168	+ 9.00	138.81	8.71	252	Q94168 potentilla anserina. pol.
sp_invertebrate:Q9XY08	+ 9.00	137.88	8.56	289	Q9XY08 strongylocentrotus pur.
sp_plant:Q9ZSK1	+ 9.00	136.61	8.36	348	Q9ZSK1 arabidopsis thaliana (n).
sp_plant:Q9XIP9	+ 9.00	136.61	8.36	348	Q9XIP9 arabidopsis thaliana (n).
sp_plant:Q9LMR3	+ 9.00	136.42	8.33	358	Q9LMR3 arabidopsis thaliana (n).
sp_mammal:Q9T7R2	+ 9.00	136.38	8.33	360	Q9T7R2 ovis aries (sheep). deca.
sp_human:Q9ULQ7	+ 9.00	135.06	8.12	437	Q9ULQ7 homo sapiens (human).
sp_human:Q9ULM6	+ 9.00	133.19	7.85	575	Q9ULM6 homo sapiens (human).
sp_human:Q9HDC7	+ 9.00	132.61	7.76	626	Q9HDC7 homo sapiens (human).
sp_human:Q9HDC7	+ 9.00	132.61	7.76	626	Q9HDC7 mus musculus (mouse).
sp_human:Q9QK78	+ 9.00	131.37	7.59	752	Q9QK78 homo sapiens (human).
sp_human:Q9QMS7	+ 9.00	130.95	7.53	799	Q9QMS7 homo sapiens (human).
sp_mammal:Q9TSP2	+ 9.00	130.73	7.50	826	Q9TSP2 papio anubis (olive babo).
sp_mammal:Q9O206	+ 9.00	130.62	7.48	839	Q9O206 homo sapiens (human).
sp_mammal:Q9T2N0	+ 9.00	130.62	7.48	839	Q9T2N0 pan paniscus (pygmy chim).
sp_plant:Q9LNX8	+ 9.00	129.81	7.37	945	Q9LNX8 arabidopsis thaliana (n).
sp_invertebrate:Q9V701	+ 9.00	129.75	7.36	953	Q9V701 drosophila melanogaster.
sp_plant:Q9LXZ4	+ 9.00	129.34	7.31	1012	Q9LXZ4 glycine max (soybean).
sp_plant:Q9J7J8	+ 9.00	127.79	7.10	1271	Q9J7J8 mus musculus (mouse).
sp_invertebrate:Q9VG11	+ 9.00	127.36	7.04	1354	Q9VG11 drosophila melanogaster.
sp_plant:Q9C6R1	+ 9.00	125.48	6.80	1784	Q9C6R1 arabidopsis thaliana (n).
sp_invertebrate:Q9I761	+ 9.00	123.79	6.59	2287	Q9I761 caenorhabditis elegans.
sp_invertebrate:Q90YG4	+ 8.00	127.92	115.21	77	Q90YG4 oncorhynchus mykiss (rain

sp_virus:Q98158	-	8.00	126.49	112.20	95	Q98158 kaposi's sarcoma-associ
sp_mammal:Q9ND09	-	8.00	125.75	110.66	106	Q9ND09 macaca fascicularis (
sp_mammal:Q95JK7	-	8.00	125.75	110.66	106	Q95JK7 macaca fascicularis (
sp_bacteria:Q05260	+	8.00	125.31	109.77	113	Q05260 bacillus subtilis. hy
sp_bacteria:Q9RPX2	+	8.00	125.07	109.29	117	Q9RPX2 haemophilus ducreyi.
sp_vertebrate:Q918G7	+	8.00	124.10	107.33	135	Q918G7 xenopus laevis (afri
sp_bacteriap:Q930L0	-	8.00	123.66	106.46	144	Q930L0 rhizobium meliloti (s
sp_bacteriap:Q97G02	-	8.00	123.43	106.00	149	Q97G02 clostridium acetobuty
sp_mammal:Q9MZA9	+	8.00	123.20	105.56	154	Q9MZA9 ovis aries (sheep). v
sp_bacteriap:Q91008	+	8.00	121.91	103.07	186	Q91008 pseudomonas aeruginos
sp_invertebrate:Q06095	+	8.00	121.81	102.87	189	Q06095 leishmania major. l
sp_bacteria:Q93JB5	+	8.00	121.73	102.73	191	Q93JB5 streptomyces coelicol
sp_rvirus:Q9QED8	+	8.00	121.45	102.20	199	Q9QED8 human immunodeficienc
sp_bacteria:Q9R6L4	+	8.00	121.42	102.14	200	Q9R6L4 agrobacterium tumefac
sp_rvirus:Q90NN1	+	8.00	121.35	102.01	202	Q90NN1 human immunodeficienc
sp_plant:Q04536	-	8.00	121.28	101.88	204	Q04536 arabidopsis thaliana
sp_bacteria:Q9KX09	+	8.00	121.22	101.75	206	Q9KX09 staphylococcus aureus
sp_virus:Q9JGH2	+	8.00	121.15	101.63	208	Q9JGH2 human immunodeficienc
sp_plant:Q9FP51	+	8.00	121.06	101.45	211	Q9FP51 oryza sativa (rice).
sp_bacteriap:Q67786	-	8.00	120.68	100.74	223	Q67786 aquifex aeolicus. eno
sp_bacteria:Q9KJL0	-	8.00	120.59	100.57	226	Q9KJL0 trichodesmium sp. ims
sp_plant:Q9FH99	-	8.00	120.53	100.46	228	Q9FH99 arabidopsis thaliana
sp_bacteria:Q91507	-	8.00	120.04	99.55	245	Q91507 methylobacterium exto
sp_organelle:Q99905	-	8.00	119.76	99.05	255	Q99905 urobatis halleri (hal
sp_bacteria:Q30597	+	8.00	119.68	98.90	258	Q30597 spiroplasma citri. hy
sp_bacteriap:Q92N10	-	8.00	119.53	98.62	264	Q92N10 rhizobium meliloti (s
sp_rvirus:Q9J4Q5	+	8.00	119.48	98.52	266	Q9J4Q5 human immunodeficienc
sp_plant:Q9LJ73	+	8.00	118.96	97.58	287	Q9LJ73 arabidopsis thaliana
sp_bacteriap:Q92LJ5	+	8.00	118.66	97.04	300	Q92LJ5 rhizobium meliloti (s
sp_bacteriap:Q9K986	-	8.00	118.48	96.72	308	Q9K986 bacillus halodurans.
sp_invertebrate:Q9N156	+	8.00	117.89	95.66	336	Q9N156 spiroplasma thaliana
sp_invertebrate:Q9VED2	+	8.00	117.72	95.38	344	Q9VED2 drosophila melanog
sp_rvirus:Q997S6	-	8.00	117.47	94.93	357	Q997S6 human immunodeficienc
sp_rvirus:Q12381	-	8.00	117.47	94.93	357	Q12381 bovine leukemia virus
sp_phage:Q38159	+	8.00	117.38	94.77	362	Q38159 bacteriophage t270. i
sp_rvirus:Q9W030	+	8.00	117.34	94.70	364	Q9W030 human immunodeficienc
sp_plant:Q9LXU5	+	8.00	117.21	94.47	371	Q9LXU5 arabidopsis thaliana
sp_organelle:Q36202	+	8.00	117.05	94.19	380	Q36202 urobatis concentricus
sp_bacteriap:Q07929	-	8.00	116.89	93.91	392	Q07929 rhodobacter capsulatu
sp_archaeo:Q58317	+	8.00	116.84	93.82	396	Q58317 pyrococcus horikoshii
sp_bacteriap:Q87677	+	8.00	116.77	93.70	396	Q87677 mycolactopsis mediter
sp_rvirus:Q96604	-	8.00	116.56	93.35	408	Q96604 avian infectious bron
sp_invertebrate:Q61449	+	8.00	116.46	93.17	414	Q61449 amblyomma americanu
sp_rvirus:Q941607	+	8.00	116.40	93.06	418	Q941607 human immunodeficienc
sp_bacteria:Q9XBW2	+	8.00	116.24	92.78	428	Q9XBW2 porphyromonas gingiva
sp_bacteriap:Q07373	+	8.00	116.19	92.70	431	Q07373 salmonella typhimurili
sp_rdent:Q9CXD9	+	8.00	116.00	92.38	443	Q9CXD9 mus musculus (mouse).
sp_plant:Q9LPM5	+	8.00	115.75	91.94	460	Q9LPM5 arabidopsis thaliana
sp_human:Q96L15	+	8.00	115.53	91.57	475	Q96L15 homo sapiens (human).
sp_vertebrate:Q91630	+	8.00	115.40	91.36	484	Q91630 xenopus laevis (afri
sp_vertebrate:Q91829	-	8.00	115.36	91.28	487	Q91829 xenopus laevis (afri
sp_bacteria:Q9ZB05	-	8.00	115.10	90.84	506	Q9ZB05 lactococcus lactis. p
sp_invertebrate:Q16628	-	8.00	114.81	90.36	528	Q16628 caenorhabditis eleg
sp_plant:Q93Z01	+	8.00	114.79	90.34	529	Q93Z01 arabidopsis thaliana
sp_plant:Q92229	+	8.00	114.72	90.21	535	Q92229 arabidopsis thaliana
sp_virus:Q916E5	+	8.00	114.55	89.93	548	Q916E5 human calicivirus nlv
sp_vertebrate:Q57653	-	8.00	114.41	89.69	560	Q57653 fugu rubripes (japan
sp_virus:Q9PN64	+	8.00	114.36	89.61	564	Q9PN64 spodoptera litura nuc
sp_plant:Q9PGH8	+	8.00	114.33	89.57	566	Q9PGH8 arabidopsis thaliana
sp_plant:Q04143	+	8.00	114.15	89.27	581	Q04143 silene latifolia. hyp
sp_invertebrate:Q46042	+	8.00	114.14	89.25	582	Q46042 drosophila melanog
sp_invertebrate:Q9W4Q7	+	8.00	114.14	89.25	582	Q9W4Q7 drosophila melanog
sp_plant:Q9FHL8	+	8.00	114.06	89.12	589	Q9FHL8 arabidopsis thaliana
sp_plant:Q9ARM5	+	8.00	113.92	88.89	601	Q9ARM5 arabidopsis thaliana
sp_plant:Q9ARE5	+	8.00	113.87	88.80	606	Q9ARE5 capsella rubella. hyp
sp_plant:Q9ZWC6	+	8.00	113.86	88.78	607	Q9ZWC6 arabidopsis thaliana
sp_plant:Q24437	+	8.00	113.80	88.69	612	Q24437 oryza longistaminata
sp_bacteriap:Q72862	+	8.00	113.64	88.42	627	Q72862 synechocystis sp. (st

sp_virus:Q86486	+	8.00	113.54	88.26	636	i	Q86486	rinderpest virus, f. prot	sp_invertebrate:Q9VXU0	+	8.00	106.88	78.02	1690	i	Q9vXu0	drosophila melanog
sp_plant:Q9LZY0	+	8.00	113.41	88.05	648	i	Q9LZY0	arabidopsis thaliana (mc	sp_invertebrate:Q16039	+	8.00	106.81	77.91	1708	i	Q16039	drosophila melanog
sp_human:Q96M50	+	8.00	113.37	87.92	656	i	Q96M50	homo sapiens (human). c	sp_vertbrate:Q9YXK0	+	8.00	106.50	77.46	1788	i	Q9YXK0	orochromis aureus (chick
sp_bacteriap:Q99U8	-	8.00	113.24	87.78	664	i	Q99U8	staphylococcus aureus (s	sp_vertbrate:Q98U10	-	8.00	105.37	75.87	2108	i	Q98U10	gallus gallus (chick
sp_plant:Q9LHNS	+	8.00	113.93	87.55	678	i	Q9LHNS	arabidopsis thaliana (mc	sp_virus:Q9DGT6	+	8.00	104.72	74.96	2321	i	Q9DGT6	turkey herpesvirus.
sp_plant:Q93ZW6	+	8.00	112.93	87.28	695	i	Q93ZW6	arabidopsis thaliana (mc	sp_plant:Q9DMZ3	-	8.00	103.89	73.81	2621	i	Q9DMZ3	arabidopsis thaliana
sp_plant:Q950027	+	8.00	112.69	86.89	720	i	Q950027	lycopersicon hirsutum, h	sp_human:Q9SGE8	-	8.00	103.79	73.68	2658	i	Q9SGE8	arabidopsis thaliana
sp_plant:Q92UH7	+	8.00	112.48	86.54	743	i	Q92UH7	arabidopsis thaliana (mc	sp_human:Q13782	+	7.00	115.61	1.5e+03	29	i	Q13782	homo sapiens (human).
sp_plant:Q92S81	+	8.00	112.25	86.18	768	i	Q92S81	lycopersicon esculentum	sp_invertebrate:Q25433	+	7.00	114.73	1.5e+03	33	i	Q25433	molgula eltrina, mem
sp_plant:Q9LFG1	+	8.00	112.12	85.97	783	i	Q9LFG1	arabidopsis thaliana (mc	sp_organelle:Q4378	+	7.00	113.25	1.4e+03	41	i	Q4378	hordeum vulgare (barle
sp_plant:Q9LST9	+	8.00	112.11	85.96	784	i	Q9LST9	arabidopsis thaliana (mc	sp_bacteriap:Q9A383	+	7.00	113.25	1.4e+03	41	i	Q9A383	caulobacter crescent
sp_plant:Q9NA83	+	8.00	112.10	85.93	786	i	Q9NA83	arabidopsis thaliana (mc	sp_bacteriap:Q92HT9	+	7.00	113.25	1.4e+03	41	i	Q92HT9	rickettsia conorii, 50
sp_plant:Q9S146	-	8.00	112.07	85.89	789	i	Q9S146	arabidopsis thaliana (mc	sp_human:Q13497	+	7.00	113.08	1.4e+03	42	i	Q13497	homo sapiens (human).
sp_human:Q9Y2C9	+	8.00	112.01	85.80	796	i	Q9Y2C9	homo sapiens (human). t	sp_todent:Q9ERA2	+	7.00	112.32	1.4e+03	47	i	Q9ERA2	microtus arvalis (comm
sp_plant:Q950025	+	8.00	111.92	85.66	806	i	Q950025	lycopersicon hirsutum, c	sp_plant:Q8566	+	7.00	110.11	1.3e+03	65	i	Q8566	catharanthus roseus (r
sp_plant:Q9C815	+	8.00	111.66	85.24	838	i	Q9C815	arabidopsis thaliana (mc	sp_bacteriap:Q9F6S9	+	7.00	109.32	1.3e+03	73	i	Q9F6S9	helicobacter pylori (c
sp_virus:Q9S491	+	8.00	111.55	85.07	851	i	Q9S491	bovine leukemia virus (b	sp_bacteriap:Q9F6S8	+	7.00	109.32	1.3e+03	73	i	Q9F6S8	helicobacter pylori (c
sp_virus:Q9S494	+	8.00	111.55	85.06	852	i	Q9S494	bovine leukemia virus (b	sp_bacteriap:Q47834	+	7.00	109.32	1.3e+03	73	i	Q47834	enterococcus hirae, ar
sp_plant:Q92S82	+	8.00	111.54	85.05	853	i	Q92S82	lycopersicon esculentum	sp_plant:Q9CAC6	+	7.00	109.32	1.3e+03	73	i	Q9CAC6	arabidopsis thaliana
sp_plant:Q950024	+	8.00	111.52	85.02	855	i	Q950024	lycopersicon hirsutum, a	sp_human:Q9NU40	+	7.00	108.87	1.3e+03	78	i	Q9NU40	homo sapiens (human).
sp_plant:Q92S83	+	8.00	111.52	85.02	855	i	Q92S83	lycopersicon esculentum	sp_plant:Q9AG48	+	7.00	108.36	1.3e+03	84	i	Q9AG48	poncirus trifoliata, p
sp_plant:Q950026	+	8.00	111.47	84.94	862	i	Q950026	lycopersicon hirsutum, h	sp_virus:Q95349	+	7.00	108.04	1.3e+03	88	i	Q95349	autographa californica
sp_plant:Q940235	+	8.00	111.46	84.92	863	i	Q940235	lycopersicon pimpinellif	sp_bacteriap:Q9F8U2	-	7.00	107.97	1.3e+03	89	i	Q9F8U2	streptomyces rishirien
sp_plant:Q946D6	+	8.00	111.46	84.92	863	i	Q946D6	lycopersicon pimpinellif	sp_bacteriap:Q9AMX5	-	7.00	107.67	1.3e+03	93	i	Q9AMX5	bradyrhizobium japonic
sp_plant:Q95WT8	+	8.00	111.44	84.90	865	i	Q95WT8	hordeum vulgare (barley)	sp_mammal:Q9NO42	-	7.00	107.59	1.3e+03	94	i	Q9NO42	macaca fascicularis (c
sp_plant:Q92S79	+	8.00	111.44	84.90	865	i	Q92S79	lycopersicon esculentum	sp_bacteriap:Q9CCD5	-	7.00	107.59	1.3e+03	94	i	Q9CCD5	mycobacterium leprae.
sp_plant:Q950022	+	8.00	111.44	84.90	865	i	Q950022	lycopersicon pimpinellif	sp_invertebrate:Q9NFH0	+	7.00	107.31	1.3e+03	98	i	Q9NFH0	plasmodium falciparu
sp_human:Q9H465	+	8.00	111.44	84.89	866	i	Q9H465	homo sapiens (human). ba	sp_organelle:Q35843	+	7.00	107.31	1.3e+03	98	i	Q35843	sceloporus grammicus.
sp_plant:Q950021	+	8.00	111.44	84.89	866	i	Q950021	lycopersicon pimpinellif	sp_virus:Q90177	-	7.00	107.11	1.3e+03	101	i	Q90177	human immunodeficien
sp_plant:Q9FJX6	+	8.00	111.18	84.49	899	i	Q9FJX6	arabidopsis thaliana (mc	sp_archaeap:Q9VE29	-	7.00	106.97	1.3e+03	103	i	Q9VE29	aeropyrum pernix, hyp
sp_invertebrate:Q9VP24	+	8.00	111.16	84.45	902	i	Q9VP24	drosophila melanogast	sp_fungi:Q12111	-	7.00	106.91	1.3e+03	104	i	Q12111	saccharomyces cerevisi
sp_plant:Q9C905	+	8.00	111.10	84.36	910	i	Q9C905	arabidopsis thaliana (mc	sp_human:Q9HAZ6	+	7.00	106.91	1.3e+03	104	i	Q9HAZ6	homo sapiens (human).
sp_invertebrate:Q26872	+	8.00	111.06	84.30	915	i	Q26872	trypanosoma cruzi, t	sp_organelle:Q9MF8C	+	7.00	106.71	1.3e+03	107	i	Q9MF8C	beta vulgaris (sugar
sp_plant:Q9AT03	+	8.00	110.98	84.17	926	i	Q9AT03	tritium aestivum (wheat	sp_plant:Q9LXK1	+	7.00	106.71	1.3e+03	107	i	Q9LXK1	arabidopsis thaliana
sp_human:Q9H8E3	+	8.00	110.96	84.15	928	i	Q9H8E3	homo sapiens (human). co	sp_virus:Q93489	-	7.00	106.71	1.3e+03	107	i	Q93489	bovine ephemeral feve
sp_human:Q9NYD1	+	8.00	110.80	83.89	951	i	Q9NYD1	homo sapiens (human). g	sp_bacteriap:Q92GV8	-	7.00	106.71	1.3e+03	107	i	Q92GV8	rickettsia conorii, h
sp_human:Q9XB81	+	8.00	110.80	83.89	951	i	Q9XB81	homo sapiens (human). g	sp_virus:Q98963	+	7.00	106.59	1.3e+03	109	i	Q98963	human herpesvirus 6.
sp_plant:Q9CE39	+	8.00	110.75	83.81	958	i	Q9CE39	hordeum vulgare (barley)	sp_bacteriap:Q86982	-	7.00	106.52	1.3e+03	110	i	Q86982	mycobacterium tubercu
sp_plant:Q9C6A8	+	8.00	110.70	83.74	965	i	Q9C6A8	arabidopsis thaliana (mc	sp_human:Q9AXV0	-	7.00	106.52	1.3e+03	110	i	Q9AXV0	homo sapiens (human).
sp_human:Q13025	+	8.00	110.68	83.70	968	i	Q13025	homo sapiens (human). au	sp_todent:Q9C217	+	7.00	106.52	1.3e+03	110	i	Q9C217	mus musculus (mouse).
sp_bacteriap:Q9KL53	+	8.00	110.59	83.56	981	i	Q9KL53	vibrio cholerae, chemota	sp_virus:Q90623	+	7.00	106.52	1.3e+03	110	i	Q90623	simian herpes b virus
sp_plant:Q9M6A7	+	8.00	110.54	83.50	987	i	Q9M6A7	glycine max (soybean). i	sp_virus:Q9RTS0	+	7.00	106.52	1.3e+03	110	i	Q9RTS0	simian herpes b virus
sp_todent:Q9J3J0	+	8.00	110.46	83.39	997	i	Q9J3J0	mus musculus (mouse). ce	sp_phage:Q94MY3	+	7.00	106.46	1.3e+03	111	i	Q94MY3	haemophilus influenza
sp_human:Q9NQ36	+	8.00	110.46	83.37	999	i	Q9NQ36	homo sapiens (human). ce	sp_plant:Q9SFF8	+	7.00	106.34	1.3e+03	113	i	Q9SFF8	arabidopsis thaliana
sp_plant:Q940640	+	8.00	110.29	83.10	1025	i	Q940640	oryza sativa (rice), an	sp_plant:Q9MLW6	-	7.00	106.34	1.3e+03	113	i	Q9MLW6	arabidopsis thaliana
sp_plant:Q9SE97	+	8.00	110.12	82.84	1051	i	Q9SE97	arabidopsis thaliana (m	sp_todent:Q9QV52	-	7.00	106.34	1.3e+03	113	i	Q9QV52	rattus sp. e-selectin
sp_plant:Q9FGW1	+	8.00	110.09	82.80	1055	i	Q9FGW1	arabidopsis thaliana (m	sp_archaeap:Q58699	+	7.00	106.34	1.3e+03	113	i	Q58699	pyrococcus horikoshil
sp_mammal:Q77690	+	8.00	109.98	82.63	1072	i	Q77690	bos taurus (bovine). o	sp_bacteriap:Q51922	-	7.00	106.22	1.2e+03	115	i	Q51922	streptomyces coelicol
sp_plant:Q9LPU24	+	8.00	109.68	82.18	1120	i	Q9LPU24	arabidopsis thaliana (m	sp_plant:Q9AWM8	+	7.00	106.22	1.2e+03	115	i	Q9AWM8	mus musculus (mouse).
sp_invertebrate:Q9VFG3	-	8.00	109.68	82.17	1121	i	Q9VFG3	drosophila melanogast	sp_todent:Q9DAB6	+	7.00	106.16	1.2e+03	116	i	Q9DAB6	mus musculus (mouse).
sp_plant:Q9SUB9	+	8.00	109.54	81.97	1143	i	Q9SUB9	arabidopsis thaliana (m	sp_fungi:Q01614	+	7.00	106.05	1.2e+03	118	i	Q01614	pneumocystis carinii.
sp_bacteriap:Q97LNB	-	8.00	109.40	81.75	1167	i	Q97LNB	clostridium acetobutyl	sp_invertebrate:Q9BII0	+	7.00	106.05	1.2e+03	118	i	Q9BII0	anopheles gambiae (c
sp_human:Q9NSE1	+	8.00	109.35	81.67	1176	i	Q9NSE1	homo sapiens (human). i	sp_bacteriap:Q84774	-	7.00	105.99	1.2e+03	119	i	Q84774	chlamydia trachomatis
sp_plant:Q92V43	+	8.00	109.29	81.58	1187	i	Q92V43	arabidopsis thaliana (m	sp_bacteriap:Q92815	+	7.00	105.99	1.2e+03	119	i	Q92815	chlamydia pneumoniae
sp_human:Q9P2J5	+	8.00	109.14	81.36	1212	i	Q9P2J5	homo sapiens (human). k	sp_bacteriap:Q9KD06	-	7.00	105.93	1.2e+03	120	i	Q9KD06	bacillus halodurans.
sp_plant:Q9PH15	+	8.00	109.13	81.34	1214	i	Q9PH15	arabidopsis thaliana (m	sp_archaeap:Q58480	+	7.00	105.93	1.2e+03	120	i	Q58480	pyrococcus horikoshil
sp_human:Q13826	+	8.00	109.13	81.34	1215	i	Q13826	homo sapiens (human). a	sp_bacteriap:Q983W0	+	7.00	105.82	1.2e+03	122	i	Q983W0	rhizobium loti (mesor
sp_invertebrate:Q9VJQ0	+	8.00	109.12	81.33	1216	i	Q9VJQ0	drosophila melanogast	sp_vertbrate:Q983M5	-	7.00	105.71	1.2e+03	124	i	Q983M5	agkistrodon acutus (h
sp_plant:Q9SWH3	+	8.00	109.12	81.33	1216	i	Q9SWH3	chlamydomonas reinhard	sp_bacteriap:Q9RYN3	-	7.00	105.60	1.2e+03	126	i	Q9RYN3	streptomyces coelicol
sp_archaeap:Q58836	-	8.00	109.07	81.24	1226	i	Q58836	mechanococcus jannasch	sp_plant:Q9FZT5	+	7.00	105.60	1.2e+03	126	i	Q9FZT5	arabidopsis thaliana
sp_plant:Q9R123	+	8.00	109.01	81.16	1236	i	Q9R123	arabidopsis thaliana (m	sp_todent:Q9CRN7	+	7.00	105.54	1.2e+03	127	i	Q9CRN7	mus musculus (mouse).
sp_plant:Q94107	+	8.00	108.82	80.87	1271	i	Q94107	lycopersicon esculentum	sp_bacteriap:Q92T84	-	7.00	105.54	1.2e+03	127	i	Q92T84	rhizobium melliloti (s
sp_plant:Q9LGS0	+	8.00	108.71	80.71	1292	i	Q9LGS0	oryza sativa (rice), es	sp_archaeap:Q9YA23	-	7.00	105.54	1.2e+03	127	i	Q9YA23	aeropyrum pernix, hyp
sp_human:Q9JFP8	+	8.00	108.46	80.33	1341	i	Q9JFP8	homo sapiens (human). t	sp_fungi:Q9US83	-	7.00	105.49	1.2e+03	128	i	Q9US83	schizosaccharomyces p
sp_bacteriap:Q50177	+	8.00	108.16	79.89	1401	i	Q50177	mycobacterium leprae, p	sp_human:Q96LJ4	+	7.00	105.49	1.2e+03	128	i	Q96LJ4	homo sapiens (human).
sp_virus:Q67631	+	8.00	108.08	79.77	1417	i	Q67631	gallid herpesvirus 1, i	sp_fungi:Q07790	+	7.00	105.33	1.2e+03	131	i	Q07790	saccharomyces cerevisi
sp_plant:Q9LRR5	+	8.00	108.08	79.77	1417	i	Q9LRR5	arabidopsis thaliana (m	sp_invertebrate:Q95T33	+	7.00	105.28	1.2e+03	131	i	Q95T33	drosophila melanog
sp_invertebrate:Q9NK96	+	8.00	107.83	79.40	1471	i	Q9NK96	drosophila melanogast	sp_virus:Q9IGX4	+	7.00	105.18	1.2e+03	134	i	Q9IGX4	tobacco geminivirus.
sp_invertebrate:Q9V																	

sp_plant:Q39649	7.00	104.83	1.2e+03	141	Q39649 cucurbita sp. prepro2s a	sp_bacteria:Q93ER5	7.00	103.09	1.2e+03	182	Q39e5 photorhabdus luminesc
sp_bacteria:Q93F15	7.00	104.78	1.2e+03	142	Q93f15 pseudobacter rodentium. h	sp_bacteria:Q93BMO	7.00	103.06	1.2e+03	183	Q93bm0 acinetobacter sp. put
sp_virus:Q69341	7.00	104.78	1.2e+03	142	Q69341 pseudobacter rodentium. h	sp_rodent:Q93K14	7.00	102.98	1.2e+03	185	Q93k14 mus musculus (mouse).
sp_fungi:Q95Y7L6	7.00	104.74	1.2e+03	143	Q95y7l6 schizosaccharomyces pombe	sp_bacteria:Q51021	7.00	102.98	1.2e+03	185	Q51021 rhodococcus erythrops
sp_plant:Q959X5	7.00	104.74	1.2e+03	143	Q959x5 arabidopsis thaliana (mc)	sp_bacteria:Q53147	7.00	102.95	1.2e+03	186	Q53147 rhodococcus erythrops
sp_plant:Q9XHW7	7.00	104.74	1.2e+03	143	Q9xhw7 oryza sativa (rice). put	sp_bacteria:Q91573	7.00	102.91	1.2e+03	187	Q91573 streptococcus pyogenes
sp_bacteria:Q94ZY2	7.00	104.69	1.2e+03	144	Q94zy2 pseudomonas aeruginosa. h	sp_plant:Q94D45	7.00	102.91	1.2e+03	187	Q94d45 oryza sativa (rice).
sp_archeap:Q74077	7.00	104.69	1.2e+03	144	Q74077 pseudomonas aeruginosa. h	sp_bacteria:Q930683	7.00	102.87	1.2e+03	188	Q930683 borrelia burgdorferi
sp_bacteria:Q9KNS56	7.00	104.64	1.2e+03	145	Q9kns56 vibrio cholerae. hypote	sp_invertebrate:Q9TVU3	7.00	102.84	1.2e+03	189	Q9tvu3 caenorhabditis eleg
sp_vertebrate:Q9DFF8	7.00	104.64	1.2e+03	145	Q9dff8 vibrio cholerae. hypote	sp_bacteria:Q9WZT0	7.00	102.80	1.2e+03	190	Q9wzt0 thermotoga maritima.
sp_vertebrate:Q9DG31	7.00	104.60	1.2e+03	146	Q9dg31 agkistrodon halys pallas	sp_bacteria:Q9W0895	7.00	102.76	1.2e+03	191	Q9w0895 borrelia garinii. sur
sp_archeap:Q58420	7.00	104.60	1.2e+03	146	Q58420 pyrococcus horikoshii. h	sp_human:Q9C062	7.00	102.76	1.2e+03	191	Q9c062 homo sapiens (human).
sp_archeap:Q9V2D5	7.00	104.60	1.2e+03	146	Q9v2d5 pyrococcus horikoshii. h	sp_plant:Q91I20	7.00	102.76	1.2e+03	191	Q91i20 oryza sativa (rice).
sp_human:Q9BQY1	7.00	104.55	1.2e+03	147	Q9bqy1 homo sapiens (human). d3	sp_archeap:Q9YFW9	7.00	102.76	1.2e+03	191	Q9yfw9 aeropyrum pernix. hyp
sp_mammal:Q9TU08	7.00	104.55	1.2e+03	147	Q9tu08 ovine aries (sheep). d3	sp_invertebrate:Q902245	7.00	102.73	1.2e+03	192	Q902245 caenorhabditis eleg
sp_invertebrate:Q9XUG4	7.00	104.50	1.2e+03	148	Q9xug4 caenorhabditis elegans	sp_plant:Q95DR4	7.00	102.73	1.2e+03	192	Q95dr4 thuja plicata (giant
sp_virus:Q9WRR6	7.00	104.46	1.2e+03	149	Q9wrr6 macaca mulatta rhadinov	sp_plant:Q95Y78	7.00	102.73	1.2e+03	192	Q95y78 arabidopsis thaliana
sp_archeap:Q9YF22	7.00	104.46	1.2e+03	149	Q9yf22 aeropyrum pernix. hypote	sp_rvirus:Q91Q07	7.00	102.73	1.2e+03	192	Q91q07 human immunodeficienc
sp_archeap:Q976T3	7.00	104.46	1.2e+03	149	Q976t3 sulfobolus tokodaii. hyp	sp_rvirus:Q91Q06	7.00	102.73	1.2e+03	192	Q91q06 human immunodeficienc
sp_organella:Q9MQV6	7.00	104.41	1.2e+03	150	Q9mqv6 simulium congreanarum. h	sp_fungi:Q12064	7.00	102.69	1.2e+03	193	Q12064 saccharomyces cerevis
sp_invertebrate:Q91600	7.00	104.37	1.2e+03	151	Q91600 chironomus thummi (mic	sp_bacteria:Q52612	7.00	102.66	1.2e+03	194	Q52612 borrelia burgdorferi
sp_invertebrate:Q91592	7.00	104.37	1.2e+03	151	Q91592 chironomus pallidivitt	sp_bacteria:Q92379	7.00	102.66	1.2e+03	194	Q92379 borrelia burgdorferi
sp_invertebrate:Q91594	7.00	104.37	1.2e+03	151	Q91594 chironomus pallidivitt	sp_bacteria:Q92H98	7.00	102.66	1.2e+03	194	Q92h98 borrelia burgdorferi
sp_plant:Q93573	7.00	104.32	1.2e+03	152	Q93573 solanum tuberosum (potat	sp_plant:Q65863	7.00	102.66	1.2e+03	194	Q65863 pinus elliottii (slas
sp_plant:Q93VW7	7.00	104.28	1.2e+03	153	Q93vw7 oryza sativa (rice). ojl	sp_bacteria:Q9ZH90	7.00	102.62	1.2e+03	195	Q9zh90 borrelia burgdorferi
sp_fungi:Q008540	7.00	104.23	1.2e+03	154	Q008540 saccharomyces cerevisia	sp_bacteria:Q50902	7.00	102.62	1.2e+03	195	Q50902 myxococcus xanthus. a
sp_mammal:Q9XS38	7.00	104.23	1.2e+03	154	Q9xs38 papio hamadryas (hamadry	sp_rodent:Q88324	7.00	102.59	1.2e+03	196	Q88324 mus musculus (mouse).
sp_virus:Q9BMO	7.00	104.23	1.2e+03	154	Q9bmo0 amseta moorei entomopox	sp_vertebrate:Q90259	7.00	102.59	1.2e+03	196	Q90259 brachydanio rerio (ze
sp_invertebrate:Q962T5	7.00	104.19	1.2e+03	155	Q962t5 spodioptra frugiperda	sp_bacteria:Q9PK45	7.00	102.59	1.2e+03	196	Q9pk45 chlamydia muridarum.
sp_rodent:Q9CV25	7.00	104.19	1.2e+03	155	Q9cv25 mus musculus (mouse). 23	sp_bacteria:Q52326	7.00	102.55	1.2e+03	197	Q52326 synechococcus sp. (st
sp_bacteria:Q9CGG06	7.00	104.14	1.2e+03	156	Q9cgg06 lactococcus lactis (subs	sp_bacteria:Q50029	7.00	102.55	1.2e+03	197	Q50029 mycobacterium leprae.
sp_rodent:Q9P776	7.00	104.10	1.2e+03	157	Q9p7t6 mus musculus (mouse). 22	sp_plant:Q9M4Y3	7.00	102.55	1.2e+03	197	Q9m4y3 mesembryanthemum crys
sp_bacteria:Q93S76	7.00	104.06	1.2e+03	158	Q93st6 chlorobium tepidum. bchf	sp_plant:Q93265	7.00	102.55	1.2e+03	197	Q93265 arabidopsis thaliana
sp_invertebrate:Q9NA77	7.00	104.06	1.2e+03	158	Q9na77 caenorhabditis elegans	sp_archeap:Q9YEV5	7.00	102.55	1.2e+03	197	Q9yev5 aeropyrum pernix. hyp
sp_bacteria:Q9ZNJ9	7.00	104.01	1.2e+03	159	Q9znj9 clostridium histolyticum	sp_plant:Q92R37	7.00	102.52	1.2e+03	198	Q92r37 arabidopsis thaliana
sp_plant:Q9M635	7.00	104.01	1.2e+03	159	Q9m636 glycine max (soybean). h	sp_bacteria:Q9XAP3	7.00	102.49	1.2e+03	199	Q9xap3 streptomyces coelicol
sp_plant:Q9C7P7	7.00	103.93	1.2e+03	161	Q9c7p7 arabidopsis thaliana (mc	sp_invertebrate:Q9VTA9	7.00	102.49	1.2e+03	199	Q9vta9 drosophila melanoga
sp_invertebrate:Q9NBH8	7.00	103.89	1.2e+03	162	Q9nbh8 chironomus thummi (mic	sp_invertebrate:Q44230	7.00	102.49	1.2e+03	199	Q44230 anthicidaris crassi
sp_invertebrate:Q9NBH7	7.00	103.89	1.2e+03	162	Q9nbh7 chironomus thummi (mic	sp_vertebrate:Q9PTE4	7.00	102.49	1.2e+03	199	Q9pte4 brachydanio rerio (ze
sp_bacteria:Q9PML3	7.00	103.89	1.2e+03	162	Q9pml3 campylobacter jejuni. pc	sp_fungi:Q9US77	7.00	102.42	1.2e+03	201	Q9us77 schizosaccharomyces p
sp_bacteria:Q92H51	7.00	103.89	1.2e+03	162	Q92h51 rickettsia conorii. hpc	sp_invertebrate:Q45632	7.00	102.42	1.2e+03	201	Q45632 caenorhabditis eleg
sp_invertebrate:Q01885	7.00	103.84	1.2e+03	163	Q01885 caenorhabditis elegans	sp_bacteria:Q93KP9	7.00	102.38	1.2e+03	202	Q93kp9 bacillus pasteurii. u
sp_rodent:Q97875	7.00	103.84	1.2e+03	163	Q97875 rattus norvegicus (rat).	sp_invertebrate:Q9W1M5	7.00	102.38	1.2e+03	202	Q9w1m5 drosophila melanoga
sp_virus:Q89338	7.00	103.84	1.2e+03	163	Q89338 african swine fever viru	sp_invertebrate:Q9VWM2	7.00	102.38	1.2e+03	202	Q9vwm2 drosophila melanoga
sp_bacteria:Q9CNA3	7.00	103.84	1.2e+03	163	Q9cna3 pasteurella multocida. h	sp_invertebrate:Q9VWM1	7.00	102.38	1.2e+03	202	Q9vwm1 drosophila melanoga
sp_plant:Q9FT76	7.00	103.80	1.2e+03	164	Q9ft76 arabidopsis thaliana (mc	sp_bacteria:Q92B62	7.00	102.38	1.2e+03	202	Q92b62 listeria innocua. lin
sp_human:Q9P1P2	7.00	103.72	1.2e+03	166	Q9p1p2 homo sapiens (human). pr	sp_bacteria:Q9X7L5	7.00	102.35	1.2e+03	203	Q9x7l5 rhizobium meliloti (s
sp_invertebrate:Q9NGV6	7.00	103.72	1.2e+03	166	Q9ngv6 leishmania infantum. e	sp_bacteria:Q9ZDB0	7.00	102.35	1.2e+03	203	Q9zdb0 rickettsia prowazekii
sp_bacteria:Q52171	7.00	103.68	1.2e+03	167	Q52171 salmonella berta. orfix.	sp_archeap:Q9H160	7.00	102.28	1.2e+03	205	Q9h160 thermoplasma acidophi
sp_invertebrate:Q02370	7.00	103.68	1.2e+03	167	Q02370 chironomus thummi thum	sp_invertebrate:Q9VGG2	7.00	102.25	1.2e+03	206	Q9vvg2 drosophila melanoga
sp_invertebrate:Q9N7D0	7.00	103.68	1.2e+03	167	Q9n7d0 leishmania major. poss	sp_invertebrate:Q61793	7.00	102.25	1.2e+03	206	Q61793 caenorhabditis eleg
sp_plant:Q82293	7.00	103.68	1.2e+03	167	Q82293 arabidopsis thaliana (mc	sp_bacteria:Q31254	7.00	102.22	1.2e+03	207	Q31254 anabaena sp. cytochro
sp_bacteria:Q9JY55	7.00	103.68	1.2e+03	167	Q9jy55 neisseria meningitidis (	sp_invertebrate:Q9W1F0	7.00	102.22	1.2e+03	207	Q9w1f0 drosophila melanoga
sp_invertebrate:Q9VLJ9	7.00	103.64	1.2e+03	168	Q9vlj9 drosophila melanogaste	sp_fungi:Q9USB7	7.00	102.12	1.2e+03	210	Q9usb7 schizosaccharomyces p
sp_plant:Q9SE91	7.00	103.60	1.2e+03	169	Q9se91 brassica oleracea (cauli	sp_invertebrate:Q9VGA9	7.00	102.12	1.2e+03	210	Q9vga9 drosophila melanoga
sp_invertebrate:Q9BIL7	7.00	103.56	1.2e+03	170	Q9bil7 giardia lamblia (giard	sp_organella:Q9G914	7.00	102.09	1.2e+03	211	Q9g914 ochromonas danica. or
sp_plant:Q9N0H1	7.00	103.56	1.2e+03	170	Q9n0h1 phaeooides dalli (dall	sp_plant:Q9LPE8	7.00	102.09	1.2e+03	211	Q9lpe8 arabidopsis thaliana
sp_plant:Q95I87	7.00	103.56	1.2e+03	170	Q95i87 arabidopsis thaliana (mc	sp_invertebrate:Q9U8F1	7.00	102.05	1.2e+03	212	Q9u8f1 schistosoma japonic
sp_bacteria:Q9JY37	7.00	103.48	1.2e+03	172	Q9jy37 neisseria meningitidis (	sp_virus:Q91SY8	7.00	102.05	1.2e+03	212	Q91sy8 hepatitis b virus. pr
sp_invertebrate:Q969B1	7.00	103.44	1.2e+03	173	Q969b1 giardia lamblia (giard	sp_fungi:Q03387	7.00	101.99	1.2e+03	214	Q03387 saccharomyces cerevis
sp_human:Q9NRB6	7.00	103.40	1.2e+03	174	Q9nrb6 homo sapiens (human). mu	sp_plant:Q9FV23	7.00	101.99	1.2e+03	214	Q9fv23 oryza sativa (rice).
sp_invertebrate:Q961D8	7.00	103.36	1.2e+03	174	Q961d8 drosophila melanogaste	sp_archeap:Q58541	7.00	101.99	1.2e+03	214	Q58541 methanococcus jannasc
sp_rodent:Q9D800	7.00	103.32	1.2e+03	176	Q9d800 mus musculus (mouse). 20	sp_fungi:Q9US44	7.00	101.96	1.2e+03	215	Q9us44 schizosaccharomyces p
sp_invertebrate:Q9C9Y8	7.00	103.28	1.2e+03	176	Q9c9y8 mus musculus (mouse). 13	sp_invertebrate:Q17578	7.00	101.96	1.2e+03	215	Q17578 caenorhabditis eleg
sp_fungi:Q9P542	7.00	103.24	1.2e+03	178	Q9p542 neurospora crassa. conse	sp_rodent:Q08904	7.00	101.93	1.2e+03	216	Q08904 mus musculus (mouse)
sp_invertebrate:Q9XWH9	7.00	103.24	1.2e+03	178	Q9xwh9 caenorhabditis elegans	sp_bacteria:Q9PCA2	7.00	101.90	1.2e+03	217	Q9pca2 streptomyces coelicol
sp_plant:Q9FP14	7.00	103.21	1.2e+03	179	Q9fp14 arabidopsis thaliana (mc	sp_plant:Q9USJ9	7.00	101.90	1.2e+03	217	Q9usj9 arabidopsis thaliana
sp_bacteria:Q47021	7.00	103.21	1.2e+03	179	Q47021 escherichia coli. 'ori'	sp_plant:Q9XJ08	7.00	101.90	1.2e+03	217	Q9xj08 arabidopsis thaliana
sp_invertebrate:Q9VCB4	7.00	103.17	1.2e+03	180	Q9vcb4 drosophila melanogaste	sp_bacteria:Q9XZ79	7.00	101.86	1.2e+03	218	Q9xz79 thermotoga maritima.
sp_bacteria:Q9APV2	7.00	103.17	1.2e+03	180	Q9apv2 pseudomonas aeruginosa.	sp_rodent:Q9DBJ3	7.00	101.83	1.2e+03	219	Q9dbj3 mus musculus (mouse).
sp_invertebrate:Q96298	7.00	103.17	1.2e+03	180	Q96298 vorticella convallaria	sp_human:Q9NW37	7.00	101.77	1.2e+03	221	Q9nw37 homo sapiens (human).
sp_plant:Q9M203	7.00	103.13	1.2e+03	181	Q9m203 arabidopsis thaliana (mc	sp_plant:Q94HL8	7.00	101.77	1.2e+03	221	Q94hl8 oryza sativa (rice).
sp_plant:Q9LGF0	7.00	103.13	1.2e+03	181	Q9lgf0 oryza sativa (rice). p04	sp_invertebrate:Q9VWM4	7.00	101.74	1.2e+03	222	Q9vwm4 drosophila melanoga

sp_vertibrate:Q91980	-	7.00	101.74	1.1e+03	222	Q919b0 xenopus laevis (african	+	7.00	100.90	1.1e+03	251	Q9pyu2 xestia c-nigrum granu
sp_bacteriap:Q72733	-	7.00	101.74	1.1e+03	222	P72733 synecocystis sp. (strai	+	7.00	100.88	1.1e+03	251	Q94190 duchesnea indica. pol
sp_bacteriap:Q9A601	-	7.00	101.74	1.1e+03	222	Q9A601 cauleobacter crescentus	+	7.00	100.88	1.1e+03	252	Q94189 duchesnea indica. pol
sp_fungi:Q07988	+	7.00	101.68	1.1e+03	224	Q07988 saccharomyces cerevisiae	+	7.00	100.88	1.1e+03	252	Q94188 duchesnea indica. pol
sp_archeap:Q9Y908	+	7.00	101.68	1.1e+03	224	Q9Y908 aeropyrum pernix. hypoth	+	7.00	100.88	1.1e+03	252	Q94184 horkelia cuneata. pol
sp_bacteriap:Q9AHL4	+	7.00	101.65	1.1e+03	225	Q9AHL4 pasteurella multocida. h	+	7.00	100.88	1.1e+03	252	Q94167 potentilla fruticosa. pol
sp_plant:Q9A3A4	+	7.00	101.65	1.1e+03	225	Q9A3A4 oryza sativa (rice). put	+	7.00	100.85	1.1e+03	253	Q99102 mus musculus (mouse).
sp_human:Q9B8G8	-	7.00	101.62	1.1e+03	226	Q9B8G8 homo sapiens (human). si	-	7.00	100.82	1.1e+03	254	Q47169 escherichia coli. out
sp_human:Q961L9	-	7.00	101.62	1.1e+03	226	Q961L9 homo sapiens (human). co	-	7.00	100.82	1.1e+03	254	Q9V3r0 drosophila melanoga
sp_bacteriap:Q935V9	-	7.00	101.59	1.1e+03	227	Q935V9 streptomyces sp. ec3. tr	-	7.00	100.82	1.1e+03	254	P70636 rattus norvegicus (ra
sp_plant:Q9F0S9	-	7.00	101.59	1.1e+03	227	Q9F0S9 arabidopsis thaliana (mc	-	7.00	100.82	1.1e+03	254	P91we2 mus musculus (mouse).
sp_bacteriap:Q9K7U1	+	7.00	101.53	1.1e+03	227	Q9K7U1 bacillus halodurans. bh3	+	7.00	100.82	1.1e+03	254	Q981c4 rhizobium loti (mesor
sp_human:Q9H6A4	+	7.00	101.53	1.1e+03	229	Q9H6A4 homo sapiens (human). co	+	7.00	100.80	1.1e+03	255	Q93gd1 desulfotomaculum geot
sp_bacteriap:Q9L1N3	+	7.00	101.50	1.1e+03	230	Q9L1N3 streptomyces coelicolor	+	7.00	100.80	1.1e+03	255	Q93gd1 desulfotomaculum geot
sp_human:Q9H3K3	+	7.00	101.50	1.1e+03	230	Q9H3K3 homo sapiens (human). my	+	7.00	100.80	1.1e+03	255	Q9665 caenorhabditis eleg
sp_invertebrate:Q9N3T3	+	7.00	101.50	1.1e+03	230	Q9N3T3 caenorhabditis elegans	+	7.00	100.80	1.1e+03	255	Q94cd0 oryza sativa (rice).
sp_plant:Q9SFE4	+	7.00	101.50	1.1e+03	230	Q9SFE4 arabidopsis thaliana (mc	+	7.00	100.74	1.1e+03	257	Q9a22 streptomyces violaceo
sp_plant:Q9A4L83	+	7.00	101.50	1.1e+03	230	Q9A4L83 horkelia cuneata. polyga	+	7.00	100.74	1.1e+03	257	Q98b97 rhizobium loti (mesor
sp_bacteriap:Q9HU29	+	7.00	101.50	1.1e+03	231	Q9HU29 pseudomonas aeruginosa.	+	7.00	100.72	1.1e+03	258	Q99p23 shigella flexneri. or
sp_bacteriap:Q9F5X4	+	7.00	101.47	1.1e+03	231	Q9F5X4 bacillus licheniformis.	+	7.00	100.69	1.1e+03	259	Q96303 streptomyces ambofaci
sp_fungi:Q74961	+	7.00	101.47	1.1e+03	231	Q74961 schizosaccharomyces pombe	+	7.00	100.69	1.1e+03	259	Q95649 streptomyces viridoch
sp_human:Q9H8T1	+	7.00	101.47	1.1e+03	231	Q9H8T1 homo sapiens (human). hy	+	7.00	100.66	1.1e+03	260	Q9r1w5 streptomyces coelicol
sp_rodent:Q9CWC5	-	7.00	101.44	1.1e+03	232	Q9CWC5 mus musculus (mouse). 90	-	7.00	100.66	1.1e+03	260	Q9adH4 streptomyces coelicol
sp_bacteriap:Q9HW17	+	7.00	101.44	1.1e+03	232	Q9HW17 pseudomonas aeruginosa.	+	7.00	100.66	1.1e+03	260	Q94386 streptomyces lividans
sp_bacteriap:Q9FMX3	+	7.00	101.44	1.1e+03	232	Q9FMX3 clostridium acetobutylic	+	7.00	100.66	1.1e+03	260	Q9u417 plodia interpunctel
sp_plant:Q9AV43	+	7.00	101.41	1.1e+03	233	Q9AV43 oryza sativa (rice). hyp	+	7.00	100.66	1.1e+03	260	Q9u416 plodia interpunctel
sp_bacteriap:Q50136	+	7.00	101.41	1.1e+03	233	Q50136 mycobacterium leprae. al	+	7.00	100.66	1.1e+03	260	Q9sw12 arabidopsis thaliana
sp_bacteriap:Q9LB88	+	7.00	101.38	1.1e+03	234	Q9LB88 polyangium cellulorum. h	+	7.00	100.64	1.1e+03	261	Q97ck9 thermoplasma volcanu
sp_invertebrate:Q9YMW3	-	7.00	101.38	1.1e+03	234	Q9YMW3 drosophila melanogaste	-	7.00	100.64	1.1e+03	261	Q9Xa33 streptomyces coelicol
sp_mammal:Q9RTJ3	+	7.00	101.38	1.1e+03	234	Q9RTJ3 equus caballus (horse).	+	7.00	100.64	1.1e+03	261	Q9vwx7 drosophila melanog
sp_virus:Q8S310	+	7.00	101.38	1.1e+03	234	Q8S310 Shope fibroma virus (str	+	7.00	100.64	1.1e+03	261	Q963598 plodia interpunctel
sp_virus:Q9Q8G0	+	7.00	101.38	1.1e+03	234	Q9Q8G0 myxoma virus (strain lat	+	7.00	100.61	1.1e+03	262	Q9av55 hordeum vulgare (barl
sp_bacteriap:Q9JLF6	+	7.00	101.35	1.1e+03	234	Q9JLF6 rhizobium loti (mesorhiz	+	7.00	100.61	1.1e+03	262	Q940h0 arabidopsis thaliana
sp_mammal:Q95JP5	+	7.00	101.35	1.1e+03	235	Q95JP5 macaca fascicularis (cra	+	7.00	100.61	1.1e+03	262	Q91m02 groundnut rosette vir
sp_bacteriap:Q54014	+	7.00	101.32	1.1e+03	236	Q54014 paracoccus denitrificans	+	7.00	100.61	1.1e+03	262	Q91584 pseudomonas aeruginos
sp_bacteriap:Q10531	+	7.00	101.29	1.1e+03	237	Q10531 mycobacterium tuberculosis	+	7.00	100.59	1.1e+03	263	Q9m6c2 neosporea crassa. hy
sp_archeap:Q26387	+	7.00	101.29	1.1e+03	237	Q26387 methanothermobacter ther	+	7.00	100.59	1.1e+03	263	Q9p9c3 arabidopsis thaliana
sp_bacteriap:Q9ZN98	+	7.00	101.27	1.1e+03	238	Q9ZN98 escherichia coli. penici	+	7.00	100.56	1.1e+03	264	Q9ae26 rhizobium leguminosar
sp_bacteriap:Q9Z027	+	7.00	101.24	1.1e+03	239	Q9Z027 rhizobium meliloti (sinc	+	7.00	100.56	1.1e+03	264	Q97n52 streptococcus pneumon
sp_plant:Q9SP59	+	7.00	101.21	1.1e+03	240	Q9SP59 populus x canescens. ara	+	7.00	100.56	1.1e+03	264	Q97779 methanothermobacter t
sp_fungi:Q02685	-	7.00	101.18	1.1e+03	241	Q02685 saccharomyces cerevisiae	-	7.00	100.53	1.1e+03	265	Q9Z401 pseudomonas putida. t
sp_plant:Q9SP08	+	7.00	101.18	1.1e+03	241	Q9SP08 arabidopsis thaliana (mc	+	7.00	100.53	1.1e+03	265	Q9U4z8 manduca sexta (toba
sp_human:Q9UG10	+	7.00	101.15	1.1e+03	242	Q9UG10 homo sapiens (human). hy	+	7.00	100.53	1.1e+03	265	Q923581 caenorhabditis eleg
sp_bacteriap:Q86207	+	7.00	101.12	1.1e+03	243	Q86207 enterococcus faecalis (hy	+	7.00	100.53	1.1e+03	265	Q9Hwv2 pseudomonas aeruginos
sp_plant:Q49965	+	7.00	101.12	1.1e+03	243	Q49965 brassica campestris (fife	+	7.00	100.51	1.1e+03	266	Q9RY20 deinococcus radiodura
sp_archeap:Q9YF02	+	7.00	101.12	1.1e+03	243	Q9YF02 aeropyrum pernix. hypoth	+	7.00	100.46	1.1e+03	266	Q97N94 streptococcus pneumonia
sp_bacteriap:Q9Z556	+	7.00	101.10	1.1e+03	244	Q9Z556 zymomonas mobilis. ria F	+	7.00	100.46	1.1e+03	268	Q9slt3 streptomyces coelicol
sp_invertebrate:Q9U462	-	7.00	101.10	1.1e+03	244	Q9U462 drosophila melanogaste	-	7.00	100.46	1.1e+03	268	Q93g19 uncultured marine bac
sp_organelle:Q94Z17	+	7.00	101.10	1.1e+03	244	Q94Z17 pyralia littoralis. se	+	7.00	100.46	1.1e+03	268	Q93ss9 homo sapiens (human).
sp_plant:Q63627	+	7.00	101.10	1.1e+03	244	Q63627 arabidopsis thaliana (mc	+	7.00	100.46	1.1e+03	268	Q99xg8 streptococcus pyogene
sp_plant:Q91SJ7	+	7.00	101.10	1.1e+03	244	Q91SJ7 arabidopsis thaliana (mc	+	7.00	100.46	1.1e+03	269	Q94bq9 arabidopsis thaliana
sp_plant:Q9LND1	+	7.00	101.10	1.1e+03	244	Q9LND1 arabidopsis thaliana (mc	+	7.00	100.41	1.1e+03	270	Q9rrv3 deinococcus radiodura
sp_plant:Q9FL17	+	7.00	101.10	1.1e+03	244	Q9FL17 arabidopsis thaliana (mc	+	7.00	100.38	1.1e+03	271	Q991g6 mus musculus (mouse).
sp_bacteriap:Q9F8U9	+	7.00	101.07	1.1e+03	245	Q9F8U9 streptomyces rishirensi	+	7.00	100.36	1.1e+03	272	Q9Xvex7 caenorhabditis eleg
sp_bacteriap:Q47168	+	7.00	101.07	1.1e+03	245	Q47168 escherichia coli. outer	+	7.00	100.36	1.1e+03	272	Q9fk56 arabidopsis thaliana
sp_organelle:Q95F05	+	7.00	101.07	1.1e+03	245	Q95F05 lunularia cruciata. puta	+	7.00	100.36	1.1e+03	272	Q980422 oryza sativa (rice).
sp_plant:Q94147	+	7.00	101.04	1.1e+03	246	Q94147 spiraea cantoniensis. po	+	7.00	100.36	1.1e+03	273	Q987z6 rhizobium loti (mesor
sp_bacteriap:Q25219	+	7.00	101.04	1.1e+03	246	Q25219 helicobacter pylori (cam	+	7.00	100.33	1.1e+03	273	Q9aee3 leptospira interrogan
sp_bacteriap:Q9ZM00	+	7.00	101.04	1.1e+03	246	Q9ZM00 helicobacter pylori j99	+	7.00	100.33	1.1e+03	273	P79192 macaca mulatta (rhesu
sp_bacteriap:Q987C1	+	7.00	101.04	1.1e+03	246	Q987C1 rhizobium loti (mesorhiz	+	7.00	100.33	1.1e+03	273	Q9a628 oryza sativa (rice).
sp_plant:Q9ARD9	+	7.00	101.01	1.1e+03	247	Q9ARD9 gossypium hirsutum (upla	+	7.00	100.33	1.1e+03	273	Q9a628 oryza sativa (rice).
sp_plant:Q84507	+	7.00	101.01	1.1e+03	247	Q84507 parametium bursaria (chld	+	7.00	100.33	1.1e+03	273	Q9a628 oryza sativa (rice).
sp_archeap:Q9HRV8	+	7.00	101.01	1.1e+03	247	Q9HRV8 halobacterium sp. (strai	+	7.00	100.31	1.1e+03	274	Q99sw2 arabidopsis thaliana
sp_bacteriap:Q9HWX9	+	7.00	100.99	1.1e+03	248	Q9HWX9 pseudomonas aeruginosa.	+	7.00	100.31	1.1e+03	274	Q47074 escherichia coli. bfp
sp_bacteriap:Q91575	+	7.00	100.96	1.1e+03	249	Q91575 streptococcus pneumoniae	+	7.00	100.31	1.1e+03	274	Q9f884 escherichia coli. per
sp_plant:Q94L85	+	7.00	100.96	1.1e+03	249	Q94L85 holodiscus microphyllus.	+	7.00	100.31	1.1e+03	274	Q9f882 escherichia coli. per
sp_plant:Q94L46	+	7.00	100.96	1.1e+03	249	Q94L46 spiraea densiflora. poly	+	7.00	100.31	1.1e+03	274	Q9ez03 escherichia coli. per
sp_plant:Q94L45	+	7.00	100.96	1.1e+03	249	Q94L45 spiraea densiflora. poly	+	7.00	100.31	1.1e+03	274	Q9ez03 escherichia coli. per
sp_invertebrate:Q02212	+	7.00	100.93	1.1e+03	250	Q02212 caenorhabditis elegans	+	7.00	100.31	1.1e+03	274	Q9f873 escherichia coli. per
sp_plant:Q9F1E9	+	7.00	100.93	1.1e+03	250	Q9F1E9 arabidopsis thaliana (mc	+	7.00	100.31	1.1e+03	274	Q9f872 escherichia coli. per
sp_plant:Q9FWL7	+	7.00	100.93	1.1e+03	250	Q9FWL7 oryza sativa (rice). put	+	7.00	100.31	1.1e+03	274	Q9f871 escherichia coli. per
sp_plant:Q9SXB5	+	7.00	100.90	1.1e+03	251	Q9SXB5 homo sapiens (human). os	+	7.00	100.31	1.1e+03	274	Q9bqy2 homo sapiens (human).
sp_human:Q67684	+	7.00	100.90	1.1e+03	251	Q67684 groundnut rosette virus	+	7.00	100.31	1.1e+03	274	Q71104 bovine adenovirus typ



sp_bacteriap:Q92JH7	-	7.00	100.31	1.1e+03	274	Q92JH7 rickettsia conorii, 2-de	sp_archeap:Q973F2	+	7.00	99.55	1.1e+03	306	Q973F2 sulfolobus tokodaii.
sp_bacteria:Q9REB4	-	7.00	100.26	1.1e+03	276	Q9REB4 escherichia coli, intimi	sp_human:Q9NMD3	+	7.00	99.53	1.1e+03	307	Q9NMD3 homo sapiens (human).
sp_invertebrate:Q16706	-	7.00	100.26	1.1e+03	276	Q16706 caenorhabditis elegans	sp_invertebrate:Q20827	-	7.00	99.53	1.1e+03	307	Q20827 caenorhabditis elegans
sp_plant:Q40750	+	7.00	100.26	1.1e+03	277	Q40750 phaseolus acutifolius (t	sp_vertebrate:Q918C0	+	7.00	99.53	1.1e+03	307	Q918C0 xenopus laevis (afri
sp_bacteria:Q93HQ9	+	7.00	100.23	1.1e+03	277	Q93HQ9 streptococcus pyogenes	sp_plant:Q82114	+	7.00	99.51	1.1e+03	308	Q82114 oryza sativa (rice).
sp_plant:Q9SR44	+	7.00	100.23	1.1e+03	277	Q9SR44 arabidopsis thaliana (mc	sp_rudent:Q9SCS9	+	7.00	99.51	1.1e+03	308	Q9SCS9 mus musculus (mouse).
sp_rudent:Q9B6A2	+	7.00	100.23	1.1e+03	277	Q9B6A2 mus musculus (mouse).	sp_bacteriap:Q9JKT7	+	7.00	99.51	1.1e+03	308	Q9JKT7 clostridium acetobuty
sp_rudent:Q9B6XS9	+	7.00	100.23	1.1e+03	277	Q9B6XS9 sulfolobus tokodaii, hy	sp_bacteriap:Q934U7	+	7.00	99.49	1.1e+03	309	Q934U7 salmonella enterica s
sp_plant:Q94AH0	+	7.00	100.21	1.1e+03	278	Q94AH0 arabidopsis thaliana (mc	sp_plant:Q9FN84	+	7.00	99.49	1.1e+03	309	Q9FN84 arabidopsis thaliana
sp_bacteriap:Q84A386	+	7.00	100.21	1.1e+03	278	Q84A386 chlamydia trachomatis, f	sp_bacteriap:Q981U5	+	7.00	99.49	1.1e+03	309	Q981U5 rhizobium loti (mesor
sp_plant:Q82508	+	7.00	100.18	1.1e+03	279	Q82508 arabidopsis thaliana (mc	sp_plant:Q9FX43	+	7.00	99.47	1.1e+03	310	Q9FX43 arabidopsis thaliana
sp_plant:Q9AW77	+	7.00	100.18	1.1e+03	279	Q9AW77 oryza sativa (rice).	sp_plant:Q9C6Z4	+	7.00	99.44	1.1e+03	311	Q9C6Z4 arabidopsis thaliana
sp_plant:Q93Z15	+	7.00	100.16	1.1e+03	280	Q93Z15 arabidopsis thaliana (mc	sp_bacteriap:Q93ES6	+	7.00	99.42	1.1e+03	312	Q93ES6 escherichia coli, gam
sp_bacteriap:P74212	+	7.00	100.16	1.1e+03	280	P74212 synechocystis sp. (strai	sp_human:Q96DP6	+	7.00	99.42	1.1e+03	312	Q96DP6 homo sapiens (human).
sp_bacteriap:Q99NM6	+	7.00	100.16	1.1e+03	280	Q99NM6 streptococcus aureus (s	sp_invertebrate:Q9V6E9	+	7.00	99.42	1.1e+03	312	Q9V6E9 drosophila melanoga
sp_invertebrate:Q9VQV0	+	7.00	100.13	1.1e+03	281	Q9VQV0 drosophila melanogaste	sp_invertebrate:Q9G6L3	-	7.00	99.42	1.1e+03	312	Q9G6L3 tribolium castaneum
sp_bacteria:Q9KW35	+	7.00	100.11	1.1e+03	282	Q9KW35 wolbachia sp. wkueyo, ou	sp_plant:Q9S174	+	7.00	99.42	1.1e+03	312	Q9S174 arabidopsis thaliana
sp_bacteriap:Q9KRY1	+	7.00	100.09	1.1e+03	283	Q9KRY1 bacillus halodurans, bho	sp_virus:Q91GK5	+	7.00	99.40	1.1e+03	313	Q91GK5 epiphyas postvittana
sp_invertebrate:Q9VPC1	+	7.00	100.06	1.1e+03	284	Q9VPC1 drosophila melanogaste	sp_bacteriap:Q9Z9F3	+	7.00	99.40	1.1e+03	313	Q9Z9F3 listeria innocua, lin
sp_plant:Q82397	+	7.00	100.06	1.1e+03	284	Q82397 arabidopsis thaliana (mc	sp_invertebrate:Q94436	-	7.00	99.38	1.1e+03	314	Q94436 chlamys nipponensis
sp_rudent:Q9CV31	+	7.00	100.06	1.1e+03	284	Q9CV31 mus musculus (mouse).	sp_plant:Q9LPI8	+	7.00	99.38	1.1e+03	314	Q9LPI8 arabidopsis thaliana
sp_archeap:Q50128	+	7.00	100.06	1.1e+03	284	Q50128 pyrococcus horikoshii, h	sp_invertebrate:Q9VMB5	+	7.00	99.36	1.1e+03	315	Q9VMB5 drosophila melanoga
sp_human:Q96W08	+	7.00	100.04	1.1e+03	285	Q96W08 homo sapiens (human).	sp_plant:Q9LNM4	+	7.00	99.36	1.1e+03	315	Q9LNM4 arabidopsis thaliana
sp_plant:Q9L281	+	7.00	100.04	1.1e+03	285	Q9L281 arabidopsis thaliana (mc	sp_bacteriap:Q9ZA18	+	7.00	99.36	1.1e+03	315	Q9ZA18 listeria innocua, lin
sp_rudent:Q9D0S4	+	7.00	100.04	1.1e+03	285	Q9D0S4 mus musculus (mouse).	sp_fungi:P78755	+	7.00	99.33	1.1e+03	316	P78755 schizosaccharomyces p
sp_human:Q93015	+	7.00	100.01	1.1e+03	286	Q93015 homo sapiens (human).	sp_bacteriap:Q98C69	+	7.00	99.33	1.1e+03	316	Q98C69 rhizobium loti (mesor
sp_plant:Q9FYH6	+	7.00	100.01	1.1e+03	286	Q9FYH6 arabidopsis thaliana (mc	sp_fungi:Q42970	+	7.00	99.31	1.1e+03	317	Q42970 schizosaccharomyces p
sp_plant:Q9LYV1	+	7.00	100.01	1.1e+03	286	Q9LYV1 arabidopsis thaliana (mc	sp_bacteriap:Q930N4	+	7.00	99.31	1.1e+03	317	Q930N4 rhizobium meliloti (s
sp_plant:Q94CM0	+	7.00	100.01	1.1e+03	286	Q94CM0 arabidopsis thaliana (mc	sp_invertebrate:Q95XN1	+	7.00	99.29	1.1e+03	318	Q95XN1 caenorhabditis eleg
sp_rudent:Q9DAE5	+	7.00	100.01	1.1e+03	286	Q9DAE5 mus musculus (mouse).	sp_plant:Q94JB7	+	7.00	99.29	1.1e+03	318	Q94JB7 oryza sativa (rice).
sp_virus:Q9DH21	+	7.00	100.01	1.1e+03	286	Q9DH21 tt virus, vp2, 12/2001	sp_rudent:Q94Z71	+	7.00	99.29	1.1e+03	318	Q94Z71 mus musculus (mouse).
sp_archeap:Q26954	+	7.00	100.01	1.1e+03	286	Q26954 methanothermobacter the	sp_bacteriap:Q55633	+	7.00	99.29	1.1e+03	318	Q55633 synechocystis sp. (st
sp_bacteria:Q48615	-	7.00	99.99	1.1e+03	287	Q48615 lactococcus lactis, abie	sp_plant:Q94K71	+	7.00	99.27	1.1e+03	319	Q94K71 arabidopsis thaliana
sp_bacteriap:Q9X099	+	7.00	99.99	1.1e+03	287	Q9X099 thermotoga maritima, con	sp_vertebrate:Q93529	+	7.00	99.27	1.1e+03	319	Q93529 xenopus laevis (afri
sp_bacteria:Q9KWR7	+	7.00	99.97	1.1e+03	288	Q9KWR7 rhodobacter capsulatus (	sp_bacteria:Q56873	+	7.00	99.23	1.1e+03	321	Q56873 versinia enterocoliti
sp_bacteria:Q9RI48	+	7.00	99.97	1.1e+03	288	Q9RI48 streptomyces coelicolor (m	sp_invertebrate:Q9BHD1	+	7.00	99.23	1.1e+03	321	Q9BHD1 leishmania major, h
sp_plant:Q9FR83	+	7.00	99.94	1.1e+03	289	Q9FR83 arabidopsis thaliana (mc	sp_bacteriap:Q968R8	+	7.00	99.23	1.1e+03	321	Q968R8 rhizobium loti (mesor
sp_plant:Q94826	+	7.00	99.94	1.1e+03	289	Q94826 oryza sativa (rice).	sp_archeap:Q9HLG3	+	7.00	99.21	1.1e+03	322	Q9HLG3 thermoplasma acidophi
sp_bacteriap:Q66723	+	7.00	99.94	1.1e+03	289	Q66723 aqifex aeolicus, hypot	sp_plant:Q82398	+	7.00	99.21	1.1e+03	322	Q82398 arabidopsis thaliana
sp_archeap:Q965YM4	+	7.00	99.94	1.1e+03	289	Q965YM4 sulfolobus tokodaii, hy	sp_plant:Q9FV11	+	7.00	99.21	1.1e+03	322	Q9FV11 vigna unguiculata (co
sp_invertebrate:Q9U394	+	7.00	99.92	1.1e+03	290	Q9U394 caenorhabditis elegans	sp_plant:Q94BM9	-	7.00	99.21	1.1e+03	322	Q94BM9 arabidopsis thaliana
sp_plant:Q94AF4	+	7.00	99.92	1.1e+03	290	Q94AF4 arabidopsis thaliana (mc	sp_rudent:Q9P627	+	7.00	99.21	1.1e+03	322	Q9P627 mus musculus (mouse).
sp_plant:Q9T013	+	7.00	99.90	1.1e+03	291	Q9T013 arabidopsis thaliana (mc	sp_organelle:Q94WF5	+	7.00	99.19	1.1e+03	323	Q94WF5 schindleria pieteschma
sp_rudent:Q9D902	+	7.00	99.87	1.1e+03	292	Q9D902 mus musculus (mouse).	sp_organelle:Q94WF4	+	7.00	99.19	1.1e+03	323	Q94WF4 schindleria pieteschma
sp_rudent:Q9D338	+	7.00	99.87	1.1e+03	292	Q9D338 mus musculus (mouse).	sp_plant:Q94AW7	+	7.00	99.19	1.1e+03	323	Q94AW7 oryza sativa (rice).
sp_rudent:Q9CWC4	+	7.00	99.87	1.1e+03	292	Q9CWC4 mus musculus (mouse).	sp_organelle:Q94WG5	+	7.00	99.16	1.1e+03	324	Q94WG5 coryphopterus persona
sp_plant:Q9LXR3	+	7.00	99.85	1.1e+03	293	Q9LXR3 arabidopsis thaliana (mc	sp_organelle:Q94WG4	+	7.00	99.16	1.1e+03	324	Q94WG4 coryphopterus hyalino
sp_bacteriap:Q99ZV7	+	7.00	99.85	1.1e+03	293	Q99ZV7 streptococcus pyogenes	sp_invertebrate:Q9VFK8	+	7.00	99.12	1.1e+03	326	Q9VFK8 drosophila melanoga
sp_bacteria:Q9K3H7	-	7.00	99.80	1.1e+03	295	Q9K3H7 streptomyces coelicolor.	sp_invertebrate:Q9U3W5	+	7.00	99.12	1.1e+03	326	Q9U3W5 drosophila melanoga
sp_bacteriap:Q9HWN6	-	7.00	99.78	1.1e+03	296	Q9HWN6 pseudomonas aeruginosa.	sp_organelle:Q9T240	+	7.00	99.12	1.1e+03	326	Q9T240 phytophthora infestan
sp_bacteriap:Q92X07	+	7.00	99.78	1.1e+03	296	Q92X07 rhizobium meliloti (sind	sp_plant:Q9SU72	+	7.00	99.12	1.1e+03	326	Q9SU72 arabidopsis thaliana
sp_bacteria:Q98084	+	7.00	99.71	1.1e+03	299	Q98084 enterococcus faecalis (s	sp_virus:Q96508	+	7.00	99.12	1.1e+03	326	Q96508 arabidopsis thaliana
sp_human:Q96LD1	+	7.00	99.71	1.1e+03	299	Q96LD1 homo sapiens (human).	sp_virus:Q9DPB4	+	7.00	99.12	1.1e+03	326	Q9DPB4 avian orthoreovirus.
sp_invertebrate:Q9UAQ5	+	7.00	99.71	1.1e+03	299	Q9UAQ5 caenorhabditis elegans	sp_plant:Q9SPC7	+	7.00	99.10	1.1e+03	327	Q9SPC7 arabidopsis thaliana
sp_plant:Q944Z4	+	7.00	99.69	1.1e+03	300	Q944Z4 arabidopsis thaliana (mc	sp_plant:Q93695	+	7.00	99.10	1.1e+03	327	Q93695 actinidopsis thaliana
sp_vertebrate:Q90ZD3	+	7.00	99.69	1.1e+03	300	Q90ZD3 oreochromis aureus (isra	sp_plant:Q48846	+	7.00	99.10	1.1e+03	327	Q48846 arabidopsis thaliana
sp_invertebrate:Q95216	-	7.00	99.67	1.1e+03	301	Q95216 trypanosoma brucei, ri	sp_bacteriap:Q9ZGN6	+	7.00	99.08	1.1e+03	328	Q9ZGN6 staphylococcus aureus
sp_plant:Q9LR34	+	7.00	99.67	1.1e+03	301	Q9LR34 arabidopsis thaliana (mc	sp_organelle:Q9XQB0	+	7.00	99.08	1.1e+03	328	Q9XQB0 phaseolus aureus (mun
sp_bacteria:Q98Q63	+	7.00	99.67	1.1e+03	301	Q98Q63 arabidopsis thaliana (mc	sp_vertebrate:Q9OX18	+	7.00	99.08	1.1e+03	328	Q9OX18 gallus gallus (chicke
sp_human:Q9BQ82	+	7.00	99.64	1.1e+03	302	Q9BQ82 homo sapiens (human).	sp_bacteriap:Q24954	+	7.00	99.08	1.1e+03	328	Q24954 helicobacter pylori (
sp_human:Q9NXX9	+	7.00	99.64	1.1e+03	302	Q9NXX9 homo sapiens (human).	sp_bacteriap:Q9ZMT8	-	7.00	99.08	1.1e+03	328	Q9ZMT8 helicobacter pylori (
sp_bacteriap:Q92RD2	+	7.00	99.64	1.1e+03	302	Q92RD2 rhizobium meliloti (sind	sp_bacteriap:Q99UA0	+	7.00	99.08	1.1e+03	328	Q99UA0 staphylococcus aureus
sp_archeap:Q59406	+	7.00	99.64	1.1e+03	302	Q59406 pyrococcus horikoshii, h	sp_bacteriap:Q99TY8	+	7.00	99.06	1.1e+03	329	Q99TY8 corynebacterium equi
sp_archeap:Q9V1H2	+	7.00	99.64	1.1e+03	302	Q9V1H2 pyrococcus abyssi, ribos	sp_plant:Q04326	+	7.00	99.06	1.1e+03	329	Q04326 arabidopsis thaliana
sp_human:Q60526	+	7.00	99.62	1.1e+03	303	Q60526 homo sapiens (human).	sp_virus:Q9WHT6	+	7.00	99.06	1.1e+03	329	Q9WHT6 bluetongue virus 4. o
sp_human:Q9UNE7	+	7.00	99.62	1.1e+03	303	Q9UNE7 homo sapiens (human).	sp_plant:Q9FL10	+	7.00	99.04	1.1e+03	330	Q9FL10 arabidopsis thaliana
sp_human:Q969U2	+	7.00	99.62	1.1e+03	303	Q969U2 homo sapiens (human).	sp_plant:Q94DN1	+	7.00	99.04	1.1e+03	330	Q94DN1 oryza sativa (rice).
sp_plant:Q94904	+	7.00	99.62	1.1e+03	303	Q94904 oryza sativa (rice).	sp_bacteriap:Q51564	+	7.00	99.04	1.1e+03	330	Q51564 borrelia burgdorferi
sp_rudent:Q923W7	+	7.00	99.62	1.1e+03	303	Q923W7 mus musculus (mouse).	sp_bacteriap:Q9CKU2	+	7.00	99.02	1.1e+03	331	Q9CKU2 pasteurella multocida
sp_rudent:Q93UD1	+	7.00	99.60	1.1e+03	304	Q93UD1 mus musculus (mouse).	sp_bacteriap:Q82937	-	7.00	99.00	1.1e+03	332	Q82937 escherichia coli o157
sp_rudent:Q9DCJ0	+	7.00	99.60	1.1e+03	304	Q9DCJ0 mus musculus (mouse).	sp_plant:Q82438	+	7.00	99.00	1.1e+03	332	Q82438 daucus carota (carrot
sp_bacteria:Q9RDP5	+	7.00	99.58	1.1e+03	305	Q9RDP5 streptomyces coelicolor.	sp_plant:Q94CB4	+	7.00	99.00	1.1e+03	332	Q94CB4 arabidopsis thaliana
sp_plant:Q9S265	+	7.00	99.55	1.1e+03	306	Q9S265 arabidopsis thaliana (mc	sp_invertebrate:Q20103	+	7.00	98.96	1.1e+03	324	Q20103 caenorhabditis eleg

sp_mhc:077979	+	7.00	98.94	1.1e+03	335	077979	rattus norvegicus (rat).	sp_plant:023397	+	7.00	98.22	1.1e+03	372	023397	arabidopsis thaliana
sp_vertibrate:09YGV5	+	7.00	98.94	1.1e+03	335	09YGV5	gallus gallus (chicken)	sp_bacteriap:09C104	+	7.00	98.22	1.1e+03	372	09C104	lactococcus lactis (s
sp_vertibrate:09YGH1	+	7.00	98.94	1.1e+03	335	09YGH1	gallus gallus (chicken)	sp_bacteriap:09P035	+	7.00	98.20	1.1e+03	373	09P035	homo sapiens (human)
sp_vertibrate:09PWR4	+	7.00	98.94	1.1e+03	335	09PWR4	gallus gallus (chicken)	sp_human:09H6B4	+	7.00	98.20	1.1e+03	373	09H6B4	homo sapiens (human)
sp_bacteriap:09AB24	+	7.00	98.94	1.1e+03	336	09AB24	caulobacter crescentus	sp_rodent:092055	+	7.00	98.20	1.1e+03	373	092055	mus musculus (mouse)
sp_plant:0948M5	+	7.00	98.90	1.1e+03	337	0948M5	oryza sativa (rice)	sp_bacteriap:099T32	-	7.00	98.19	1.1e+03	374	099T32	staphylococcus aureus
sp_bacteriap:09KWB6	+	7.00	98.86	1.1e+03	339	09KWB6	agrobacterium rhizogenes	sp_rodent:091VH1	+	7.00	98.17	1.1e+03	375	091VH1	mus musculus (mouse)
sp_invertebrate:022959	-	7.00	98.86	1.1e+03	339	022959	caenorhabditis elegans	sp_bacteriap:093P87	+	7.00	98.15	1.1e+03	376	093P87	microscilla (sp. prel.
sp_bacteriap:09KXZ70	-	7.00	98.84	1.1e+03	340	09KXZ70	streptomyces coelicolor	sp_human:09H738	+	7.00	98.15	1.1e+03	376	09H738	homo sapiens (human)
sp_bacteriap:09RV59	-	7.00	98.84	1.1e+03	340	09RV59	deinococcus radiodurans	sp_plant:09LPX9	+	7.00	98.15	1.1e+03	376	09LPX9	arabidopsis thaliana
sp_plant:09FWY3	+	7.00	98.82	1.1e+03	341	09FWY3	arabidopsis thaliana (mc	sp_archaeap:097VD4	-	7.00	98.15	1.1e+03	376	097VD4	sulfolobus solfatarica
sp_bacteriap:0875D5	+	7.00	98.80	1.1e+03	342	0875D5	campylobacter jejuni. he	sp_bacteriap:095983	+	7.00	98.13	1.1e+03	377	095983	neisseria gonorrhoeae
sp_bacteriap:09K5D3	+	7.00	98.80	1.1e+03	342	09K5D3	campylobacter jejuni. he	sp_fungi:090012	+	7.00	98.13	1.1e+03	377	090012	aspergillus aculeatus
sp_plant:0941B1	+	7.00	98.80	1.1e+03	342	0941B1	oryza sativa (rice)	sp_bacteriap:09JXX4	+	7.00	98.13	1.1e+03	377	09JXX4	neisseria meningitidis
sp_bacteriap:09PNG1	+	7.00	98.80	1.1e+03	342	09PNG1	campylobacter jejuni. pu	sp_bacteriap:09JTV26	+	7.00	98.13	1.1e+03	377	09JTV26	neisseria meningitidis
sp_bacteriap:09G5U3	+	7.00	98.78	1.1e+03	343	09G5U3	escherichia coli o157:h7	sp_mammal:09XS10	+	7.00	98.11	1.1e+03	378	09XS10	bos indicus (zebu)
sp_fungi:087179	+	7.00	98.76	1.1e+03	344	087179	schizosaccharomyces pombe	sp_plant:09FSF8	-	7.00	98.11	1.1e+03	378	09FSF8	nicotiana tabacum (co
sp_plant:09SJ03	+	7.00	98.76	1.1e+03	344	09SJ03	arabidopsis thaliana (mc	sp_virus:056768	-	7.00	98.11	1.1e+03	378	056768	human cytomegalovirus
sp_rodent:09JHR5	+	7.00	98.76	1.1e+03	344	09JHR5	mus musculus (mouse)	sp_organelle:035782	-	7.00	98.11	1.1e+03	379	035782	sorghum bicolor (sorg
sp_bacteriap:09RYS6	+	7.00	98.74	1.1e+03	345	09RYS6	deinococcus radiodurans	sp_organelle:09T9B0	+	7.00	98.10	1.1e+03	379	09T9B0	reduncula fulvofufla (
sp_organelle:078893	+	7.00	98.74	1.1e+03	345	078893	anolis coelestinus. nadi	sp_vertibrate:09DG48	+	7.00	98.10	1.1e+03	379	09DG48	xenopus laevis (afric
sp_rodent:09DA09	+	7.00	98.74	1.1e+03	345	09DA09	mus musculus (mouse)	sp_invertebrate:0965P8	-	7.00	98.08	1.1e+03	380	0965P8	caenorhabditis eleg
sp_bacteriap:09K8V2	+	7.00	98.74	1.1e+03	345	09K8V2	bacillus halodurans. n-a	sp_organelle:09T3S9	+	7.00	98.08	1.1e+03	380	09T3S9	brachyromys betsiloe
sp_bacteriap:09J512	+	7.00	98.74	1.1e+03	345	09J512	pseudomonas aeruginosa	sp_organelle:09T3F5	+	7.00	98.08	1.1e+03	380	09T3F5	brachyromys betsiloe
sp_bacteriap:09K5D3	+	7.00	98.74	1.1e+03	345	09K5D3	paracoccus pantotrophus	sp_plant:09SNQ2	+	7.00	98.08	1.1e+03	380	09SNQ2	oryza sativa (rice)
sp_bacteriap:09K5D3	+	7.00	98.74	1.1e+03	345	09K5D3	paracoccus pantotrophus	sp_rodent:061076	+	7.00	98.08	1.1e+03	380	061076	mus musculus (mouse)
sp_mammal:095J78	+	7.00	98.68	1.1e+03	348	095J78	macaca fascicularis (cre	sp_bacteriap:053592	+	7.00	98.06	1.1e+03	381	053592	streptomyces avermiti
sp_organelle:09TD42	+	7.00	98.68	1.1e+03	348	09TD42	riovulus frenatus. nadi	sp_organelle:035327	+	7.00	98.06	1.1e+03	381	035327	galactomycetes avirmiti
sp_organelle:09G0N8	+	7.00	98.68	1.1e+03	348	09G0N8	heterotaxis niloticus. nadi	sp_organelle:098H70	+	7.00	98.06	1.1e+03	381	098H70	rhizobium loti (mesor
sp_bacteriap:09ZJA5	+	7.00	98.68	1.1e+03	348	09ZJA5	ricketsia conorii. hypc	sp_invertebrate:09VY77	+	7.00	98.04	1.1e+03	382	09VY77	drosophila melanoga
sp_plant:09LSU6	+	7.00	98.66	1.1e+03	349	09LSU6	arabidopsis thaliana (mc	sp_organelle:035781	+	7.00	98.04	1.1e+03	382	035781	sorghum bicolor (sorg
sp_archaeap:09HR20	+	7.00	98.66	1.1e+03	349	09HR20	halobacterium sp. (stria	sp_plant:09BRT2	+	7.00	98.02	1.1e+03	383	09BRT2	quillardia theta (cry
sp_organelle:09HLH3	+	7.00	98.66	1.1e+03	349	09HLH3	thermoplasma acidophilum	sp_fungi:042663	+	7.00	98.02	1.1e+03	383	042663	schizosaccharomyces p
sp_human:09H6J6	-	7.00	98.66	1.1e+03	350	09H6J6	homo sapiens (human)	sp_fungi:042664	+	7.00	98.02	1.1e+03	383	042664	schizosaccharomyces p
sp_bacteriap:093FS8	+	7.00	98.62	1.1e+03	351	093FS8	sulfate-reducing bacteri	sp_invertebrate:09VDZ5	-	7.00	98.02	1.1e+03	383	09VDZ5	drosophila melanoga
sp_bacteriap:097Z26	+	7.00	98.62	1.1e+03	351	097Z26	sulfolobus solfataricus	sp_plant:09MA75	+	7.00	98.02	1.1e+03	383	09MA75	arabidopsis thaliana
sp_human:060938	+	7.00	98.60	1.1e+03	352	060938	homo sapiens (human)	sp_plant:004508	+	7.00	98.02	1.1e+03	383	004508	arabidopsis thaliana
sp_mammal:062702	+	7.00	98.60	1.1e+03	352	062702	bos taurus (bovine)	sp_vertibrate:09DEC3	+	7.00	98.02	1.1e+03	383	09DEC3	xenopus laevis (afric
sp_organelle:09GAI3	+	7.00	98.60	1.1e+03	352	09GAI3	cryptotermes austrinus	sp_bacteriap:091322	+	7.00	98.01	1.1e+03	384	091322	chlamydomophila caviae
sp_organelle:09GAI5	+	7.00	98.60	1.1e+03	352	09GAI5	cryptotermes queenslandi	sp_vertibrate:091322	+	7.00	98.01	1.1e+03	384	091322	chlamydomophila caviae
sp_human:09UJX9	+	7.00	98.58	1.1e+03	353	09UJX9	homo sapiens (human)	sp_bacteriap:09RY14	+	7.00	98.01	1.1e+03	384	09RY14	anaerococcus radiodura
sp_invertebrate:09W81	+	7.00	98.58	1.1e+03	353	09W81	drosophila melanogaster	sp_bacteriap:09SGV3	+	7.00	97.99	1.1e+03	385	09SGV3	arabidopsis thaliana
sp_organelle:09GAI9	+	7.00	98.58	1.1e+03	353	09GAI9	cryptotermes domesticus	sp_bacteriap:09UAU7	+	7.00	97.97	1.1e+03	386	09UAU7	corynebacterium equi
sp_rodent:09QZ29	-	7.00	98.58	1.1e+03	353	09QZ29	rattus norvegicus (rat)	sp_fungi:001759	+	7.00	97.97	1.1e+03	386	001759	pneumocystis carinii
sp_vertibrate:09DE66	+	7.00	98.58	1.1e+03	353	09DE66	coturnix coturnix japoni	sp_invertebrate:09BKW7	+	7.00	97.97	1.1e+03	386	09BKW7	caenorhabditis eleg
sp_vertibrate:042235	+	7.00	98.58	1.1e+03	353	042235	gallus gallus (chicken)	sp_plant:094HJ7	+	7.00	97.97	1.1e+03	386	094HJ7	oryza sativa (rice)
sp_bacteriap:097MG0	+	7.00	98.56	1.1e+03	354	097MG0	clostridium acetobutylic	sp_fungi:074747	+	7.00	97.95	1.1e+03	386	074747	schizosaccharomyces p
sp_archaeap:09Y909	+	7.00	98.56	1.1e+03	354	09Y909	aeropyrum pernix. hypoth	sp_invertebrate:022253	+	7.00	97.95	1.1e+03	387	022253	caenorhabditis eleg
sp_invertebrate:0917854	-	7.00	98.54	1.1e+03	355	0917854	caenorhabditis elegans	sp_invertebrate:023459	+	7.00	97.95	1.1e+03	387	023459	caenorhabditis eleg
sp_bacteriap:09I155	+	7.00	98.52	1.1e+03	356	09I155	streptomyces coelicolor	sp_plant:022313	+	7.00	97.95	1.1e+03	387	022313	lycopersicon esculent
sp_plant:09RJ18	-	7.00	98.50	1.1e+03	357	09RJ18	arabidopsis thaliana (mc	sp_bacteriap:09KRW8	+	7.00	97.95	1.1e+03	387	09KRW8	vibrio cholerae. hypo
sp_plant:09IPL4	+	7.00	98.45	1.1e+03	360	09IPL4	arabidopsis thaliana (mc	sp_invertebrate:09BHF3	-	7.00	97.94	1.1e+03	388	09BHF3	leishmania major. p
sp_plant:0945R1	+	7.00	98.45	1.1e+03	360	0945R1	medicago sativa (alfalfa	sp_invertebrate:09SU18	+	7.00	97.92	1.1e+03	389	09SU18	drosophila melanoga
sp_rodent:092105	+	7.00	98.43	1.1e+03	361	092105	mus musculus (mouse)	sp_plant:09FN88	+	7.00	97.92	1.1e+03	389	09FN88	arabidopsis thaliana
sp_human:096712	+	7.00	98.41	1.1e+03	362	096712	homo sapiens (human)	sp_plant:094HS3	+	7.00	97.92	1.1e+03	389	094HS3	oryza sativa (rice)
sp_invertebrate:091840	-	7.00	98.41	1.1e+03	362	091840	caenorhabditis elegans	sp_archaeap:09HIY9	-	7.00	97.90	1.1e+03	390	09HIY9	thermoplasma acidophi
sp_invertebrate:09VWV1	+	7.00	98.39	1.1e+03	363	09VWV1	drosophila melanogaster	sp_invertebrate:097747	-	7.00	97.88	1.1e+03	391	097747	caenorhabditis eleg
sp_plant:09C879	+	7.00	98.37	1.1e+03	363	09C879	arabidopsis thaliana (mc	sp_bacteriap:092M08	+	7.00	97.88	1.1e+03	391	092M08	rhizobium meliloti (s
sp_bacteriap:09CBC2	+	7.00	98.37	1.1e+03	364	09CBC2	mycobacterium leprae. 57	sp_plant:09M7Q5	+	7.00	97.87	1.1e+03	392	09M7Q5	arabidopsis thaliana
sp_plant:094AF2	+	7.00	98.35	1.1e+03	365	094AF2	arabidopsis thaliana (mc	sp_bacteriap:099X76	+	7.00	97.87	1.1e+03	392	099X76	staphylococcus aureus
sp_plant:09M670	+	7.00	98.33	1.1e+03	366	09M670	arabidopsis thaliana (mc	sp_bacteriap:09LAZ3	+	7.00	97.85	1.1e+03	393	09LAZ3	streptococcus pneumonia
sp_plant:09C933	+	7.00	98.33	1.1e+03	366	09C933	arabidopsis thaliana (mc	sp_invertebrate:018315	-	7.00	97.85	1.1e+03	393	018315	caenorhabditis eleg
sp_fungi:059847	+	7.00	98.32	1.1e+03	367	059847	aspergillus oryzae. hapi	sp_invertebrate:095TM4	-	7.00	97.85	1.1e+03	393	095TM4	drosophila melanoga
sp_bacteriap:093QC6	+	7.00	98.30	1.1e+03	368	093QC6	corynebacterium glutamic	sp_organelle:09MF43	+	7.00	97.85	1.1e+03	393	09MF43	beta vulgaris (sugar
sp_invertebrate:09BHE1	-	7.00	98.30	1.1e+03	368	09BHE1	leishmania major. prob	sp_bacteriap:085461	+	7.00	97.85	1.1e+03	393	085461	pasteurella multocida
sp_organelle:0921433	+	7.00	98.30	1.1e+03	368	0921433	taxus baccata (english y	sp_invertebrate:09VXF3	+	7.00	97.83	1.1e+03	394	09VXF3	drosophila melanoga
sp_organelle:09C2K5	+	7.00	98.30	1.1e+03	368	09C2K5	takakia lepidiozooides. r	sp_plant:09C654	-	7.00	97.83	1.1e+03	394	09C654	arabidopsis thaliana
sp_plant:09W1M4	+	7.00	98.30	1.1e+03	368	09W1M4	arabidopsis thaliana (mc	sp_archaeap:097Y05	+	7.00	97.83	1.1e+03	394	097Y05	sulfolobus solfataric
sp_human:09NQ47	+	7.00	98.26	1.1e+03	370	09NQ47	homo sapiens (human)	sp_bacteriap:09L9AY2	+	7.00	97.81	1.1e+03	395	09L9AY2	streptococcus pneumonia
sp_plant:09FHH3	+	7.00	98.24	1.1e+03	371	09FHH3	arabidopsis thaliana (mc	sp_human:09NVD0	+	7.00	97.81	1.1e+03	395	09NVD0	homo sapiens (human)
sp_plant:09AUN5	+	7.00	98.24	1.1e+03	371	09AUN5	oryza sativa (rice)	sp_bacteriap:09BRT8	+	7.00	97.81	1.1e+03	395	09BRT8	homo sapiens (human)
sp_vertibrate:09W676	+	7.00	98.24	1.1e+03	371	09W676	gallus gallus (chicken)								

sp_plant:O48581	+	7.00	97.81	1.1e+03	395	! O48581 arabidopsis thaliana (md	sp_archaeap:Q9V245	+	7.00	97.28	1.1e+03	427	! Q9V245 pyrococcus abyssi. pm
sp_human:Q9NQ55	+	7.00	97.80	1.1e+03	396	! Q9NQ55 homo sapiens (human). in	sp_bacteria:Q9APH0	-	7.00	97.27	1.1e+03	428	! Q9APH0 leptospira interrogans
sp_rhodent:Q92232	+	7.00	97.78	1.1e+03	397	! Q92232 mus musculus (mouse). un	sp_bacteria:Q97050	-	7.00	97.25	1.1e+03	429	! P7050 acetobacter pasteuria
sp_archaeap:Q97428	-	7.00	97.78	1.1e+03	397	! Q97428 sulfobolus tokodaii. put	sp_plant:O65661	+	7.00	97.25	1.1e+03	429	! O65661 arabidopsis thaliana
sp_bacteria:Q54293	-	7.00	97.75	1.1e+03	399	! Q54293 streptomyces aeruginosus	sp_rhodent:Q9V4D2	+	7.00	97.25	1.1e+03	429	! Q9V4D2 mus musculus (mouse).
sp_bacteria:Q915Q3	+	7.00	97.75	1.1e+03	399	! Q915Q3 pseudomonas aeruginosa.	sp_virus:Q9VXK3	+	7.00	97.25	1.1e+03	429	! Q9VXK3 melanoplus sanguinipes
sp_bacteria:Q9XDP7	+	7.00	97.73	1.1e+03	400	! Q9XDP7 acinetobacter calcoaceti	sp_archaeap:Q9V967	+	7.00	97.25	1.1e+03	429	! Q9V967 pyrococcus abyssi. ph
sp_invertebrate:Q9N3W0	-	7.00	97.71	1.1e+03	401	! Q9N3W0 caenorhabditis elegans	sp_fungi:Q08729	+	7.00	97.24	1.1e+03	430	! Q08729 saccharomyces cerevis
sp_plant:O82202	+	7.00	97.69	1.1e+03	402	! O82202 arabidopsis thaliana (mc	sp_organelle:Q9XMS1	+	7.00	97.24	1.1e+03	430	! Q9XMS1 trahymena pyriformi
sp_fungi:O74962	+	7.00	97.68	1.1e+03	403	! Q74962 schizosaccharomyces pombe	sp_plant:Q9PT53	+	7.00	97.24	1.1e+03	430	! Q9PT53 arabidopsis thaliana
sp_invertebrate:Q9VPD8	+	7.00	97.68	1.1e+03	403	! Q9VPD8 drosophila melanogaste	sp_bacteria:Q9Z5H7	+	7.00	97.24	1.1e+03	430	! Q9Z5H7 mycobacterium leprae.
sp_plant:Q9W1G6	-	7.00	97.68	1.1e+03	403	! Q9W1G6 arabidopsis thaliana (mc	sp_bacteria:Q9R2J2	+	7.00	97.22	1.1e+03	431	! Q9R2J2 shigella sonnei. pilo
sp_bacteria:Q912J6	-	7.00	97.68	1.1e+03	403	! Q912J6 pseudomonas aeruginosa.	sp_invertebrate:Q18874	-	7.00	97.22	1.1e+03	431	! Q18874 caenorhabditis eleg
sp_bacteria:Q91178	+	7.00	97.68	1.1e+03	403	! Q91178 pseudomonas aeruginosa.	sp_rhodent:Q9QXK6	-	7.00	97.20	1.1e+03	432	! Q9QXK6 mus musculus (mouse)
sp_bacteria:Q53233	+	7.00	97.66	1.1e+03	404	! Q53233 rhodobacter sphaeroides	sp_invertebrate:Q9VDZ4	-	7.00	97.19	1.1e+03	433	! Q9VDZ4 drosophila melanoga
sp_bacteria:Q51377	+	7.00	97.66	1.1e+03	404	! Q51377 rhodobacter sphaeroides	sp_archaeap:Q58056	+	7.00	97.19	1.1e+03	433	! Q58056 pyrococcus horikoshii
sp_invertebrate:Q16525	-	7.00	97.64	1.1e+03	405	! Q16525 caenorhabditis elegans	sp_plant:Q9CAX8	+	7.00	97.16	1.1e+03	435	! Q9CAX8 arabidopsis thaliana
sp_bacteria:Q55475	-	7.00	97.64	1.1e+03	405	! Q55475 synechocystis sp. (strai	sp_bacteria:Q98LB2	+	7.00	97.16	1.1e+03	435	! Q98LB2 rhizobium loti (mesor
sp_bacteria:Q9KTA2	+	7.00	97.63	1.1e+03	406	! Q9KTA2 deinococcus radiodurans.	sp_fungi:Q9VUG4	+	7.00	97.14	1.1e+03	436	! Q9VUG4 pichia farinosa (yeas
sp_fungi:Q94510	+	7.00	97.61	1.1e+03	407	! Q94510 schizosaccharomyces pombe	sp_invertebrate:O18714	+	7.00	97.14	1.1e+03	436	! O18714 leishmania donovani
sp_invertebrate:Q91AY0	+	7.00	97.59	1.1e+03	408	! Q91AY0 streptococcus pneumoniae	sp_rhodent:Q99P21	+	7.00	97.14	1.1e+03	436	! Q99P21 mus musculus (mouse).
sp_invertebrate:Q01634	-	7.00	97.59	1.1e+03	408	! Q01634 caenorhabditis elegans	sp_plant:Q93715	+	7.00	97.13	1.1e+03	437	! P93715 petunia hybrida (petu
sp_bacteria:Q93C87	-	7.00	97.58	1.1e+03	409	! Q93C87 mycobacterium paratuberc	sp_plant:Q9LSM4	+	7.00	97.13	1.1e+03	437	! Q9LSM4 arabidopsis thaliana
sp_invertebrate:Q9VN16	-	7.00	97.58	1.1e+03	409	! Q9VN16 drosophila melanogaste	sp_rhodent:Q91VX4	+	7.00	97.13	1.1e+03	437	! Q91VX4 mus musculus (mouse).
sp_plant:Q9FMQ9	-	7.00	97.58	1.1e+03	409	! Q9FMQ9 arabidopsis thaliana (mc	sp_plant:Q9LEQ0	+	7.00	97.11	1.1e+03	438	! Q9LEQ0 lycopersicon esculent
sp_plant:Q9SMW7	+	7.00	97.58	1.1e+03	409	! Q9SMW7 arachis hypogaea (peanut	sp_bacteria:Q9X1L8	+	7.00	97.11	1.1e+03	438	! Q9X1L8 thermotoga maritima.
sp_archaeap:Q97WQ4	+	7.00	97.58	1.1e+03	409	! Q97WQ4 sulfobolus solifataricus	sp_bacteria:Q05319	+	7.00	97.09	1.1e+03	439	! Q05319 neisseria gonorrhoeae
sp_bacteria:Q87838	-	7.00	97.56	1.1e+03	410	! Q87838 streptomyces coelicolor	sp_bacteria:Q55105	+	7.00	97.09	1.1e+03	439	! Q55105 streptococcus sp. mul
sp_organelle:Q9UMU0	+	7.00	97.56	1.1e+03	410	! Q9UMU0 mesostigma viride. hypot	sp_plant:Q9M2M7	+	7.00	97.09	1.1e+03	439	! Q9M2M7 arabidopsis thaliana
sp_bacteria:Q84710	+	7.00	97.56	1.1e+03	410	! Q84710 chlamydia trachomide. hypot	sp_bacteria:Q066730	+	7.00	97.09	1.1e+03	439	! Q066730 aquifex aeolicus. hyp
sp_bacteria:Q99R13	+	7.00	97.56	1.1e+03	411	! Q99R13 staphylococcus aureus (s	sp_bacteria:Q9CII4	+	7.00	97.08	1.1e+03	440	! Q9CII4 lactococcus lactis (s
sp_bacteria:Q93KX5	+	7.00	97.54	1.1e+03	411	! Q93KX5 streptomyces coelicolor	sp_archaeap:Q58359	+	7.00	97.08	1.1e+03	440	! Q58359 pyrococcus horikoshii
sp_invertebrate:Q9NEY6	+	7.00	97.53	1.1e+03	412	! Q9NEY6 caenorhabditis elegans	sp_bacteria:Q9HYD8	+	7.00	97.05	1.1e+03	442	! Q9HYD8 pseudomonas aeruginos
sp_bacteria:Q9CM40	+	7.00	97.51	1.1e+03	413	! Q9CM40 pasteurella multocida. p	sp_bacteria:Q9X365	+	7.00	97.03	1.1e+03	443	! Q9X365 bacillus anthracis. p
sp_mammal:Q28665	+	7.00	97.51	1.1e+03	413	! Q28665 oryctolagus cuniculus (z	sp_bacteria:Q9RSO3	-	7.00	97.02	1.1e+03	444	! Q9RSO3 deinococcus radiodura
sp_mammal:Q07238	+	7.00	97.51	1.1e+03	413	! Q07238 oryctolagus cuniculus (z	sp_bacteria:Q9KEH8	-	7.00	97.02	1.1e+03	444	! Q9KEH8 bacillus halodurans.
sp_plant:Q9SLM4	+	7.00	97.51	1.1e+03	413	! Q9SLM4 arabidopsis thaliana (mc	sp_invertebrate:Q22419	-	7.00	97.00	1.1e+03	445	! Q22419 caenorhabditis eleg
sp_bacteria:Q950214	+	7.00	97.49	1.1e+03	414	! Q950214 pseudomonas sp. vib120.	sp_invertebrate:Q27648	-	7.00	97.00	1.1e+03	445	! Q27648 entamoeba histolyti
sp_bacteria:Q9R2G0	-	7.00	97.48	1.1e+03	415	! Q9R2G0 streptomyces coelicolor.	sp_plant:Q9MBG7	+	7.00	97.00	1.1e+03	445	! Q9MBG7 arabidopsis thaliana
sp_bacteria:Q33471	-	7.00	97.48	1.1e+03	415	! Q33471 pseudomonas putida. styx	sp_organelle:Q05375	+	7.00	96.99	1.1e+03	446	! Q05375 actinobacillus actino
sp_bacteria:Q93U10	+	7.00	97.48	1.1e+03	415	! Q93U10 pseudomonas putida. styx	sp_organelle:Q9TB20	-	7.00	96.99	1.1e+03	446	! Q9TB20 drosophila immigrans
sp_plant:Q9M0F7	+	7.00	97.48	1.1e+03	415	! Q9M0F7 arabidopsis thaliana (mc	sp_organelle:Q9TB18	-	7.00	96.99	1.1e+03	446	! Q9TB18 drosophila albomicans
sp_rhodent:Q922H1	+	7.00	97.48	1.1e+03	415	! Q922H1 mus musculus (mouse). un	sp_organelle:Q9TB16	-	7.00	96.99	1.1e+03	446	! Q9TB16 drosophila albomicans
sp_virus:Q91TH4	+	7.00	97.48	1.1e+03	415	! Q91TH4 tupala herpesvirus. t122	sp_organelle:Q9TB15	-	7.00	96.99	1.1e+03	446	! Q9TB15 drosophila sulfurigas
sp_archaeap:Q9YB41	+	7.00	97.48	1.1e+03	415	! Q9YB41 aeopyrum pernix. hypoth	sp_organelle:Q9TB13	-	7.00	96.99	1.1e+03	446	! Q9TB13 drosophila sulfurigas
sp_bacteria:Q9R6S6	+	7.00	97.45	1.1e+03	417	! Q9R6S6 synechococcus sp. (strai	sp_organelle:Q9TB12	-	7.00	96.99	1.1e+03	446	! Q9TB12 drosophila sulfurigas
sp_human:Q96E08	+	7.00	97.45	1.1e+03	417	! Q96E08 homo sapiens (human). hy	sp_organelle:Q9TB10	-	7.00	96.99	1.1e+03	446	! Q9TB10 drosophila pallidifro
sp_invertebrate:Q01618	+	7.00	97.45	1.1e+03	417	! Q01618 caenorhabditis elegans	sp_organelle:Q9TB09	-	7.00	96.99	1.1e+03	446	! Q9TB09 drosophila putaua. na
sp_bacteria:Q96441	+	7.00	97.43	1.1e+03	418	! Q96441 rhizobium meliloti (sind	sp_organelle:Q9TB08	-	7.00	96.99	1.1e+03	446	! Q9TB08 drosophila nasuta. f.
sp_bacteria:Q976526	+	7.00	97.43	1.1e+03	418	! Q976526 escherichia coli. o418	sp_organelle:Q9T4V1	-	7.00	96.99	1.1e+03	446	! Q9T4V1 drosophila nasuta. f.
sp_bacteria:Q918C4	+	7.00	97.41	1.1e+03	419	! Q918C4 polyangium cellulorum. c	sp_organelle:Q9T4B3	-	7.00	96.99	1.1e+03	446	! Q9T4B3 drosophila kohkoa. an
sp_bacteria:Q9K1Z4	+	7.00	97.41	1.1e+03	419	! Q9K1Z4 polyangium cellulorum. c	sp_organelle:Q9WE32	-	7.00	96.99	1.1e+03	446	! Q9WE32 drosophila simulans (
sp_plant:Q9M1B6	+	7.00	97.41	1.1e+03	419	! Q9M1B6 arabidopsis thaliana (mc	sp_organelle:Q9MDP6	-	7.00	96.99	1.1e+03	446	! Q9MDP6 drosophila simulans (
sp_plant:Q00492	+	7.00	97.41	1.1e+03	419	! Q00492 arabidopsis thaliana (mc	sp_organelle:Q9MDK5	-	7.00	96.99	1.1e+03	446	! Q9MDK5 drosophila simulans (
sp_archaeap:Q974Y9	+	7.00	97.41	1.1e+03	419	! Q974Y9 sulfobolus tokodaii. hyp	sp_organelle:Q9MFP2	-	7.00	96.99	1.1e+03	446	! Q9MFP2 cochliomyia hominivor
sp_invertebrate:Q9P085	+	7.00	97.40	1.1e+03	420	! P90985 caenorhabditis elegans	sp_organelle:Q9B2I8	-	7.00	96.99	1.1e+03	446	! Q9B2I8 chrysomya chloropyga.
sp_bacteria:Q9CP21	-	7.00	97.38	1.1e+03	421	! Q9CP21 pasteurella multocida	sp_organelle:Q9B2I8	-	7.00	96.99	1.1e+03	446	! Q9B2I8 chrysomya chloropyga.
sp_bacteria:Q9KXV5	-	7.00	97.38	1.1e+03	421	! Q9KXV5 streptomyces coelicolor.	sp_organelle:Q9MGL7	-	7.00	96.99	1.1e+03	446	! Q9MGL7 drosophila sechellia
sp_invertebrate:Q9XVS8	+	7.00	97.38	1.1e+03	421	! Q9XVS8 caenorhabditis elegans	sp_plant:Q9RFGN4	+	7.00	96.99	1.1e+03	446	! Q9RFGN4 arabidopsis thaliana
sp_rhodent:Q91W88	+	7.00	97.38	1.1e+03	421	! Q91W88 mus musculus (mouse). ad	sp_bacteria:Q9AA93	+	7.00	96.99	1.1e+03	446	! Q9AA93 caulobacter crescentu
sp_bacteria:Q95471	+	7.00	97.35	1.1e+03	423	! Q95471 synechocystis sp. (strai	sp_bacteria:Q52382	+	7.00	96.97	1.1e+03	447	! Q52382 talmonia sp. u2. nap
sp_rhodent:Q97569	+	7.00	97.35	1.1e+03	423	! Q97569 rattus norvegicus (rat).	sp_bacteria:Q915D2	+	7.00	96.97	1.1e+03	447	! Q915D2 camamonas testosteron
sp_plant:Q92UY3	+	7.00	97.33	1.1e+03	424	! Q92UY3 arabidopsis thaliana (mc	sp_bacteria:Q915D2	+	7.00	96.97	1.1e+03	447	! Q915D2 camamonas testosteron
sp_plant:Q9SVK8	-	7.00	97.33	1.1e+03	424	! Q9SVK8 arabidopsis thaliana (mc	sp_bacteria:Q95564	+	7.00	96.97	1.1e+03	447	! Q95564 pseudomonas sp. ispal
sp_rhodent:Q99N63	+	7.00	97.33	1.1e+03	424	! Q99N63 mesocricetus auratus (gc	sp_fungi:Q9P7303	+	7.00	96.97	1.1e+03	447	! Q9P7303 schizosaccharomyces p
sp_plant:Q9R6K6	+	7.00	97.32	1.1e+03	425	! Q9R6K6 arabidopsis thaliana (mc	sp_invertebrate:Q9N4V5	+	7.00	96.97	1.1e+03	447	! Q9N4V5 caenorhabditis eleg
sp_bacteria:Q91AY5	+	7.00	97.30	1.1e+03	426	! Q91AY5 streptococcus pneumoniae	sp_plant:Q94BX2	+	7.00	96.97	1.1e+03	447	! Q94BX2 arabidopsis thaliana
sp_rhodent:Q9B655	+	7.00	97.30	1.1e+03	426	! Q9B655 streptococcus pneumoniae	sp_bacteria:Q988E5	+	7.00	96.97	1.1e+03	447	! Q988E5 rhizobium loti (mesor
sp_organelle:Q9CFH9	+	7.00	97.28	1.1e+03	427	! Q9CFH9 cercidiphyllum japonicum	sp_bacteria:Q9PLJ3	+	7.00	96.96	1.1e+03	448	! Q9PLJ3 chlamydia muridarum.
sp_plant:Q23855	+	7.00	97.28	1.1e+03	427	! Q23855 brassica campestris (fie	sp_bacteria:Q9K738	+	7.00	96.94	1.1e+03	449	! Q9K738 bacillus halodurans.
sp_bacteria:Q91186	+	7.00	97.28	1.1e+03	427	! Q91186 pseudomonas aeruginosa.	sp_rhodent:Q9DA19	+	7.00	96.93	1.1e+03	450	! Q9DA19 mus musculus (mouse).

sp\_bacteriap:Q9CM97 - 7.00 96.93 1.1e+03 450 ! Q9cm97 pasteurella multocida. h  
sp\_bacteria:Q45695 + 7.00 96.91 1.0e+03 451 ! Q45695 burkholderia sp. (strain  
sp\_plant:Q9AX41 + 7.00 96.91 1.0e+03 451 ! Q9ax41 oryza sativa (rice). p04  
sp\_archaeap:Q26414 + 7.00 96.90 1.0e+03 452 ! Q26414 methanothermobacter ther  
sp\_human:Q9UGK6 - 7.00 96.88 1.0e+03 453 ! Q9ugk6 homo sapiens (human). pu  
sp\_invertebrate:Q23615 + 7.00 96.88 1.0e+03 453 ! Q23615 caenorhabditis elegans  
sp\_human:Q9H8N3 - 7.00 96.87 1.0e+03 454 ! Q9h8n3 homo sapiens (human). cd  
sp\_invertebrate:Q9VFZ2 - 7.00 96.87 1.0e+03 454 ! Q9vfz2 drosophila melanogaste  
sp\_plant:Q9ZR98 + 7.00 96.87 1.0e+03 454 ! Q9zr98 arabidopsis thaliana (mc  
sp\_plant:Q65041 + 7.00 96.87 1.0e+03 454 ! Q65041 arabidopsis thaliana (mc  
sp\_plant:Q9A429 + 7.00 96.87 1.0e+03 454 ! Q9a429 arabidopsis thaliana (mc  
sp\_fungi:Q04172 + 7.00 96.84 1.0e+03 456 ! Q04172 saccharomyces cerevisiae

seq\_name: sp\_human.Q9NYK1

seq\_documentation\_block:  
ID Q9NYK1 PRELIMINARY; PRT; 1049 AA.  
AC Q9NYK1;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 7.  
GN TLR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Plrimates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RX MEDLINE=20477806; PubMed=11022119;  
RA Du X., Pollarak A., Wei Y., Beutler B.;  
RT "Three novel mammalian toll-like receptors: gene structure,  
expression, and evolution.";  
RL Eur. Cytokine Netw. 11:362-371(2000).  
DR EMBL; AF204067; AAF60188.1; -.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003592; LRR\_out.  
DR InterPro; IPR003591; LRR\_typ.  
DR InterPro; IPR000157; TIR.  
DR Pfam; PF00560; LRR; 12.  
DR Pfam; PF01463; LRRCT; 1.  
DR PRINTS; PRO0019; LEURICHRPT.  
DR SMART; SM00370; LRR; 4.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 3.  
DR SMART; SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 1049 AA; 120921 MW; 8C701E9E437F2721 CRC64;

alignment\_scores:  
Quality: 1049.00 Length: 1049  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9NYK1

Align seg 1/1 to: Q9NYK1 from: 1 to: 1049

85 ATGGTGTTCCTCAATGTGGACACGAAGAGACAAATCTTATCTCTTTTAA 134  
|||||  
1 MetValPheProMetTrpThrLeuLysArgGlnLeuLeuLeuPheAs 17  
135 CATAACTCAATTTCCAAACTCTTGGGGCTAGATGTTTCCCTAAACTC 184  
|||||  
17 nilelleLeuSerLysLeuLeuGlyAlaArgTrpPheProLysThrL 34

185 TGGCCCTGTGATGTGCTACTCTGGATGTTTCCAAAGAACCATGTGATCGTGGAC 234  
|||||  
34 euProCysaspValThrLeuaspValProLysAsnHisValIleValasp 50  
|||||  
235 TGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTTCCACAGAACAC 284  
|||||  
51 CysThrAspLysHisLeuThrGluIleProGlyGlyIleProThrAsnTh 67  
|||||  
285 CAGCAACCTCACCTCACCATTAAACACATACCAGACATCTCCCGAGCGT 334  
|||||  
67 rThrAsnLeuThrLeuThrIleAsnHisIleProAspIleSerProAlas 84  
|||||  
335 CCTTTCCACAGACTGGACCATCTCGTGTAGATCGATTTTCAGATGCAACTGT 384  
|||||  
84 erPheHisArgLeuaspHisLeuValGluIleaspPheArgCysAsnCys 100  
|||||  
385 GTACCTATTCCACTGGGGTCAAAAAACAACATGTCATCAAGAGGCTGCA 434  
|||||  
101 ValProIleProLeuGlySerLysAsnMetCysIleLysArgLeuG1 117  
|||||  
435 GATTAAACCCAGAGCTTTAGTGGACTCACTTATTTAAAATCCCTTTACC 484  
|||||  
117 nileLysProArgSerPheSerGlyLeuThrTyLeuLysSerLeuTyrl 134  
|||||  
485 TGGATGGAACACAGCTACTAGATACCGCAGGGCCTCCCGCCTAGCTTA 534  
|||||  
134 euaspGlyAsnGlnLeuLeuGluIleProGlnGlyLeuProProSerLeu 150  
|||||  
535 CAGCTTCTCAGCCCTTGAGGCCAACACATCTTTTCCATCAGAAAAAGAGAA 584  
|||||  
151 GlnLeuLeuSerLeuGluAlaAsnAsnIlePheSerIleArgLysGluAs 167  
|||||  
585 TCTAACAGAACTGGCCCAACATAGAAATACTCTACCTGGGCCCAAACTGTT 634  
|||||  
167 nLeuThrGluLeuAlaAsnIleGluIleLeuTyLeuGlyGlnAsnCyst 184  
|||||  
635 ATTATCGAAATCCTGTGTTTTCATATTCAATAGAGAAAGATGCTTTC 684  
|||||  
184 yTyrrArgAsnProCysTyrrValSerTyrrSerIleGluLysaspAlaPhe 200  
|||||  
685 CTAAACTTGACAAAGTTAAAGTGTCTCCCTGAAAGATAACAATGTAC 734  
|||||  
201 LeuAsnLeuThrLysLeuLysValLeuSerLeuLysAspAsnValTh 217  
|||||  
735 AGCGTCCCTACTGTTTGGCATCTACTTTAACAGAACTATATCTCTACA 784  
|||||  
217 rAlaValProThrValLeuProSerThrLeuThrGluLeuTyrrLeuTyra 234  
|||||  
785 ACAACATGATTCGCAAAATCCAAAGAGATGATTTTAAACCTCAACCAA 834  
|||||  
234 snAsnMetIleAlaLysIleGlnGluaspPheAsnAsnLeuAsnGln 250  
|||||  
835 TTTCAAAATTTCTGACCTAAGTGAATTTGCCCTCGTTGTTTATTAATGCCCC 884  
|||||  
251 LeuGlnIleLeuaspLeuSerGlyAsnCysProArgCysTyrrAsnAlaPr 267  
|||||  
885 ATTTCTCTTGTGGCCCGTGTAAAAATAATTTCTCCCTACAGATCCCTGTAA 934  
|||||  
267 oPheProCysAlaProCysLysAsnSerProLeuGlnIleProVala 284  
|||||  
935 ATGCTTTTGTGATGCGCTGACAGAATTTAAAGTTTACGCTCTACACAGTAAC 984  
|||||  
284 snAlaPheaspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300  
|||||  
985 TCTCTTTACAGATGTGCCCCCAAGATGTTTAAAGACATCAACAACTCCA 1034  
|||||  
301 SerLeuGlnHisValProProArgTrpPhePheLysAsnIleAsnLysLeuG1 317  
|||||  
1035 GGAACGTGATCTCTCCCAAACTTCTTGGCAAGAAATTTGGGGATGCTA 1084  
|||||  
317 nGluLeuaspLeuSerGlnAsnPheLeuAlaLysGluIleGlyaspAlaL 334



1085 AATTCTGCATTTCTCCCGACCTCATCCAAATGGATGTGCTTTTCAAT 1134  
|||||  
334 ysPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350  
1135 TTTGAACTTCAGGTCTATCGTCGTCATCTATGATCTATCACAGCAATTTTC 1184  
|||||  
351 PheGluLeuGlnValTyrArgAlaSerMetAsnLeuSerGlnAlaPheSe 367  
1185 TTCACGTGAAAAGCCGAAAATTCGCGGATCAGAGGATATGCTTTTAAAG 1234  
|||||  
367 rSerLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysG 384  
1235 AGTTGAAAAGCTTTAACTCTCGCCATTACATAATCTTCAAAATCTTGAA 1284  
|||||  
384 LuLeuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400  
1285 GTTCTTGATCTTGGCAGCTAACTTTTATAAAATTTGCTAACCTCAGCATGTT 1334  
|||||  
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417  
1335 TAAACAATTTAAAGACCTGAAAAGTCATAGATCTTTTCAGTGAATAAAATAT 1384  
|||||  
417 elysGlnPheLysArgLeuLysValIleAsnLeuSerValAsnLysIleS 434  
1385 CACCTTCAGGAGATTCAAAGTGAAGTTGGCTTCTGCTCAAATGCCAGAACT 1434  
|||||  
434 erProSerGlyAspSerSerGluValGlyPheCysSerAsnAlaArgThr 450  
1435 TCTGTAGAAGTTATGAACCCAGCTCTCGGAAACAATTTACATTATTTCAG 1484  
|||||  
451 SerValGluSerTyrGluProGlnValLeuGluGlnLeuHisTyrPheAr 467  
1485 ATATGATAGTATGCAAGGAGTTGCAGATTCAAAAACAAGAGGCTTCTT 1534  
|||||  
467 gTyrAspLysTyrAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484  
1535 TCATGTCTGTTTAATGAAAGCTCTACAAGTATGGCAGACCTTGGATCTA 1584  
484 heMetSerValAsnGluSerCysTyrLysTyrGlyGlnThrLeuAspLeu 500  
1585 AGTAAATAAGTATATTTTGTCTCAAGTCTCTGATTTTCAGCATCTTTC 1634  
501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517  
1635 TTTTCTCAATGCCTGAATCTCTCAGGAAATCTCATTTAGCCAAACTCTTA 1684  
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534  
1685 ATGGCAGTGAATTCACACCTTTAGCAGAGCTGAGATATTTGGACTTCTCC 1734  
534 snGlySerGluPheGlnProLeuAlaGluLeuArgTyrLeuAspPheSer 550  
1735 ACAACCCGCTTGATTTACTTCATTCACAGCATTTTGAAGACTTCACAA 1784  
551 AsnAsnArgLeuAspLeuLeuHisSerThrAlaPheGluGluLeuHisLy 567  
1785 ACTGGAAGTTCTGGATATAAGCAGTAATAGCCATTATTTTCAATCAGAG 1834  
567 sleuGluValleuAspIleSerSerAsnSerHisTyrPheGlnSerGluG 584  
1835 GAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTTCTGCAGAAA 1884  
584 lyleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600  
1885 CTGATGATGAACACAAATGACATCTCTTCTCCACCAGCAGGACCATGGA 1934  
601 LeuMetMetAsnAspAsnAspIleSerSerSerThrSerArgThrMetG1 617  
1935 GAGTGAAGTCTTAGAAGCTCTGGAATTCAGAGGAATACATTAGATGTTT 1984  
617 uSerGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspValL 634  
1985 TATGGAGAGAAGGTGATAACAGATACCTTACAATTTATCAAGAATCTGCTA 2034

634 euTrpArgGluGlyAspAsnArgTyrLeuGlnLeuPheLysAsnLeuLeu 650  
2035 RAATTAGAGGAATTAGACATCTCTAAAAATCCCTAAAGTTTCTTCGCTTC 2084  
651 LysLeuGluGluLeuAspIleSerLysAsnSerLeuSerPheLeuProSe 667  
2085 TGGAGTTTTTGTGATGCTATGCTCCCAAAATCTAAAGAAATCTCTCTTTGGCCA 2134  
667 rGlyValPheAspGlyMetProProAsnLeuLysAsnLeuSerLeuAlaL 684  
2135 AAAATGGGCTCAAAATCTTTCAGTTGGAAGAACTCCAGTGTCTAAAGAAC 2184  
684 ysAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
2185 CTGGAAACTTTGGACTCAGCCACACCAACTGACCAGTGTCCCTGAGAG 2234  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrThrValProGluAr 717  
2235 ATTATCCAACTGTTCCAGAAAGCTCAAGAAATCTGATTCTTAAGAATAATC 2284  
717 gleuSerAsnCysSerArgSerLeuLysAsnLeuIleLeuLysAsnAsnG 734  
2285 AAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCCTTCCAGTTGCGA 2334  
734 InIleArgSerLeuThrLysTyrPheLeuGlnAspAlaPheGlnLeuArg 750  
2335 TATCTGGATCTCAGCTCAATAAAATCCAGATGATCCAAAAGACCAGCTT 2384  
751 TyrLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
2385 CCAGAAAATGCTCTCAACAATCTGAAGATGTGCTTTTGCATCATATATC 2434  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuHisHisAsnA 784  
2435 GGTTCCTGTGCACCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2484  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpTrpValAsnHis 800  
2485 ACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGGGGCC 2534  
801 ThrGluValThrIleProTyrLeuAlaThrAspValThrCysValGlyPr 817  
2535 AGGAGCACACAAGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
817 oGlyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyrThrCysG 834  
2585 AGTTAGATCTGACTAACCTGATTTCTTCTCCTACTTTCCATATCTGTATCT 2634  
834 LuLeuAspLeuThrAsnLeuIleLeuPheSerLeuSerIleSerValSer 850  
2635 CTCCTTCTCATGCTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGT 2684  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyrPheTrpAspVa 867  
2685 GTGGTATATTTACCATTTCTCTGAAGGCCAAGATAAAGGGGTATCAGCGTC 2734  
867 ltrpTyrIleTyrHisPheCysLysAlaLysIleLysGlyTyrGlnArgL 884  
2735 TAATATCACAGACTGTTGCTATGATGCTTTTATTTGTGTATGACATAAA 2784  
884 euIleSerProAspCysCysTyrAspAlaPheIleValTyrAspThrLys 900  
2785 GACCAGCTGTGACCGAGTGGGTTTTTGGCTGAGCTGGTGGCCAAACTGGA 2834  
901 AspProAlaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuG1 917  
2835 AGACCCCAAGAGAGAACATTTAAATTTATGTCTCGAGGAAAGGACTGGT 2884  
917 uAspProArgGluLysHisPheAsnLeuCysLeuGluGluArgAspTrpL 934  
2885 TACCAGGGCAGCCAGTCTCTGGAAAACCTTTCCAGAGCATACAGCTTAGC 2934  
|||||



935 ATGCTTTTGATGCGCTGACAGAAATTTAAAGTTTACGCTACACAGTAAC 984  
|||||  
284 snAlaPheAspAlaLeuThrGluLeuLysValLeuArgLeuHisSerAsn 300  
|||||  
985 TCTCTTCAGCATGTGCCCCCAAGATGGTTTAAAGAACATCAACAAACTCCA 1034  
|||||  
301 SerLeuGlnHisValProProArgTrpPheLysAsnIleAsnLysLeuGI 317  
|||||  
1035 GGAACGGATCTGCTCCCAAACTTCTTGCCAAAGAAATTTGGGGATGCTA 1084  
|||||  
317 nGluLeuAspLeuSerGlnAsnPheLeuAlaLysGluIleGlyAspAlaL 334  
|||||  
1085 AATTTCTGATTTTCTCCCGAGCTCATCAATTTGGATCTGTCTTCAAT 1134  
|||||  
334 ysPheLeuHisPheLeuProSerLeuIleGlnLeuAspLeuSerPheAsn 350  
|||||  
1135 TTTGAACTTCAGGCTCATCGTCATCTATGAATCTATCACAAAGCATTTTC 1184  
|||||  
351 PheGluLeuGlnValIlyrArgAlaSerMetAsnLeuSerGlnAlaPheSe 367  
|||||  
1185 TTCACCTGAAAGCCTGAAATTTCTGCGGATCAGAGGATATGTCTTTAAAG 1234  
|||||  
367 rSerLeuLysSerLeuLysIleLeuArgIleArgGlyTyrValPheLysG 384  
|||||  
1235 AGTTGAAAGCTTTAACTCTCCCATTTACATAATCTTCAAAATCTTGAA 1284  
|||||  
384 luLeuLysSerPheAsnLeuSerProLeuHisAsnLeuGlnAsnLeuGlu 400  
|||||  
1285 GTTCTTTGATCTTGGCACTTAACCTTTATATAAATTTGCTTAACCTCAGCATGTT 1334  
|||||  
401 ValLeuAspLeuGlyThrAsnPheIleLysIleAlaAsnLeuSerMetPh 417  
|||||  
1335 TAAACAATTTAAAGACTGAAAGCTATAGATCTTTCAGTGAATAAATAT 1384  
|||||  
417 eLysGlnPheLysArgLeuLysValIleAspLeuSerValAsnLysIleS 434  
|||||  
1385 CACCTTCAGAGATTCAGTGAAGTTGGTCTGCTCAAAATGCCAGAACT 1434  
|||||  
434 erProSerGlyAspSerSerGluValGlyPheCysSerAsnAlaArgThr 450  
|||||  
1435 TCTGTAGAAGTTATCAACCCAGGTCCTGGAACAATTCATTTATTCAG 1484  
|||||  
451 SerValGluSerTyrGluProGlnValLeuGluGlnLeuHisTyrPheAr 467  
|||||  
1485 ATATGATAGTATGCAAGGAGTTGCAAGATTCAAAACAAAGAGCCTCTT 1534  
|||||  
467 gTyrAspLysTyrAlaArgSerCysArgPheLysAsnLysGluAlaSerP 484  
|||||  
1535 TCATGCTGTTAATGAAGCTGTACAAGTATGGCAGACCTTTGGATCTA 1584  
|||||  
484 heMetSerValAsnGluSerCysTyrLysTyrGlyGlnThrLeuAspLeu 500  
|||||  
1585 AGTAAATAAGTATATTTTGTCAAGTCTCTGATTTTCAGCATCTTTC 1634  
|||||  
501 SerLysAsnSerIlePhePheValLysSerSerAspPheGlnHisLeuSe 517  
|||||  
1635 TTTCTCCAATGCTGAATCTGCAGAAATCTCATTAGCCAAACTCTTA 1684  
|||||  
517 rPheLeuLysCysLeuAsnLeuSerGlyAsnLeuIleSerGlnThrLeuA 534  
|||||  
1685 ATGGCAGTGAATTCACACCTTTAGCAGAGCTGAGATATTTGGACTTCTCC 1734  
|||||  
534 snGlySerGluPheGlnProLeuAlaGluLeuArgTyrLeuAspPheSer 550  
|||||  
1735 AACACCGGCTTGTATTTACTCCATTCAACAGCATTTGAAGAGCTTTCACAA 1784  
|||||  
551 AsnAsnArgLeuAspLeuLeuHisSerThrAlaPheGluGluLeuHisLy 567  
|||||  
1785 ACTGGAAGTTCTGGATATAGCAGTAAATAGCCATTTATTTCAATCAGAAG 1834  
|||||  
567 sLeuGluValLeuAspIleSerSerAsnSerHisTyrPheGlnSerGluG 584  
|||||  
1835 GAATTACTCATATGCTAAACTTTTACCAAGAACCTTAAAGGTTCTGCAGAA 1884  
|||||

584 lyIleThrHisMetLeuAsnPheThrLysAsnLeuLysValLeuGlnLys 600  
|||||  
1885 CTGATCATGAACGACAATGACATCTCTCTCCACCAGCAGGACCATTGGA 1934  
|||||  
601 LeuMetMetAsnAspAsnAspIleSerSerThrSerArgThrMetGI 617  
|||||  
1935 GAGTGAAGTCTCTTAGAACTCTGGAATTCAGAGGAAATCACATTAGATGTTT 1984  
|||||  
617 uSerGluSerLeuArgThrLeuGluPheArgGlyAsnHisLeuAspValL 634  
|||||  
1985 TATGGAGAGAAAGTGATACAGATACATATTACAAATTATTCAAGATCTGCTA 2034  
|||||  
634 euTrpArgGluGlyAspAsnArgTyrLeuGlnLeuPheLysAsnLeuLeu 650  
|||||  
2035 AAATTAGAGAAATTAGACATCTCTAAAAATTCCTTAAGTTTCTTGCTTTC 2084  
|||||  
651 LysLeuGluGluLeuAspIleSerLysAsnSerLeuSerPheLeuProSe 667  
|||||  
2085 TGGAGTTTTTGTATGCTCCTCCAAATCTAAAGAAATCTCTCTTTGGCCA 2134  
|||||  
667 rGlyValPheAspGlyMetProProAsnLeuLysAsnLeuSerLeuAlaL 684  
|||||  
2135 AAAATGGGCTCAAACTCTTCAGTTGGAAGAACTCCAGTGTCTTAAAGAAC 2184  
|||||  
684 ysAsnGlyLeuLysSerPheSerTrpLysLysLeuGlnCysLeuLysAsn 700  
|||||  
2185 CTGGAACCTTTGGACCTCAGCCCAACCAACTGACCACCTGCTCCCTGAGAG 2234  
|||||  
701 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThrValProGluAr 717  
|||||  
2235 ATTATCCAACCTGTTCCAGAGCCTCAAGAAATCTGATTCTTAAAGAAATATC 2284  
|||||  
717 gLeuSerAsnCysSerArgSerHisLysAsnLeuIleLeuLysAsnAsnG 734  
|||||  
2285 AAATCAGGAGTCTGCAAGATATTTTCTACAAGATGCTTTCAGTTGGCGA 2334  
|||||  
734 InIleArgSerProThrLysTyrPheLeuGlnAspAlaPheGlnLeuArg 750  
|||||  
2335 TATCTGATCTCAGCTCAATAAAATCCAGATGATCCAAAGAACCCAGCTT 2384  
|||||  
751 TyrLeuAspLeuSerSerAsnLysIleGlnMetIleGlnLysThrSerPh 767  
|||||  
2385 CCAGAAAAATGTCTCAACAATCTGAAGATGTTGCTTTTGCATCATATAATC 2434  
|||||  
767 eProGluAsnValLeuAsnAsnLeuLysMetLeuLeuLeuHisHisAsnA 784  
|||||  
2435 GGTTCCTGTCACCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2484  
|||||  
784 rgPheLeuCysThrCysAspAlaValTrpPheValTrpTrpValAsnHis 800  
|||||  
2485 ACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTGTGGGGCC 2534  
|||||  
801 ThrGluValThrIleProTyrLeuAlaThrAspValThrCysValGlyPr 817  
|||||  
2535 AGGAGCACACAGGGCCAAAGTGTGATCTCCCTGGATCTGTACACCTGTG 2584  
|||||  
817 oGlyAlaHisLysGlyGlnSerValIleSerLeuAspLeuTyrThrCysG 834  
|||||  
2585 AGTTAGATCTGACTAACCTGATTCTGTCTCATTCTCCATATCTGATCT 2634  
|||||  
834 luLeuAspLeuThrAsnLeuIleLeuPheSerLeuSerIleSerValSer 850  
|||||  
2635 CTCTTTCTCATGTTGATGATGACAGCAAGTCACTCTATTCTGGGATGT 2684  
|||||  
851 LeuPheLeuMetValMetMetThrAlaSerHisLeuTyrPheTrpAspVa 867  
|||||  
2685 GTGGTATATTTACCATTTCTGTAAGCCAAAGATAAAGGGGTATCAGCGTC 2734  
|||||  
867 lTrpTyrIleTyrHisPheCysLysAlaLysIleLysGlyTyrGlnArgL 884  
|||||  
2735 TAATATCACCAGACTGTGCTATGATGCTTTTATTTGTTGATGACACTAAA 2784  
|||||

```
884 euIleSerProAspCysCysTyraSpAlaPheIleValTyraSpThrLys 900
2785 GACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGGTCCCAACTGGA 2834
|||||
901 AspprolaValThrGluTrpValLeuAlaGluLeuValAlaLysLeuG 917
|||||
2835 AGACCCAGAGAGAAACATTTTAAATTTATGCTCGAGGAAAGGACTGGT 2884
|||||
917 uAspproArgGluLysHisPheAsnLeuCysLeuGluGluArgAspTrpL 934
|||||
2885 TACCAGGCGCAGCTGTCTCGAAACCTTTCCAGAGCATTACAGCTTAGC 2934
|||||
934 euProGlyGlnProValLeuGluAsnLeuSerGlnSerIleGlnLeuSer 950
|||||
2935 AAAAAGACAGTGTGTGATGACAGACAACTATCCAAAGACTGAAATTT 2984
|||||
951 LysLysThrValPheValMetThrAspLysTyraLysThrGluAsnPh 967
|||||
2985 TAGATAGCATTTTACTGTGCCATCAGAGGCTCATGTGATGAAAAAGTTG 3034
|||||
967 eLysIleAlaPheTyrlLeuSerHisGlnArgLeuMetAspGluLysVala 984
|||||
3035 ATGTGATATCTGTGATTTCTGTGAGAACCTTTTCAGAGTCCAAGTTC 3084
|||||
984 spValIleIleLeuIlePheLeuGluLysProPheGlnLysSerLysPhe 1000
|||||
3085 CTCAGCTCCGAAAGGCTCTGTGGAGTCTGCTCTGAGTGGCAAC 3134
|||||
1001 LeuGlnLeuArgLysArgLeuCysGlySerValLeuGluTrpProTh 1017
|||||
3135 AAACCCGCAAGCTCACCATCTCTGGCAGTGTCTAAAGAACGCCCTGG 3184
|||||
1017 rAsnProGlnAlaHisProTyrlPheTrpGlnCysLeuLysAsnAlaLeuA 1034
|||||
3185 CCACAGACAATCATGTGGCTATAGTCAGCTGTTCAAGGAACGGTC 3231
|||||
1034 laThrAspAsnHisValAlaTyrlSerGlnValPheLysGluThrVal 1049
|||||
```

seq\_name: sp\_rodent:Q92311

seq\_documentation\_block:  
ID Q92311 PRELIMINARY; PRT; 1050 AA.

AC Q92311;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 7.

GN TLR7.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;  
RT "Molecular cloning of murine Toll-Like-Receptor 7";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035889; AAK62676.1; -;  
KW Receptor.  
SQ SEQUENCE 1050 AA; 121836 MW; 495B75DEE849D8EE CRC64;

alignment\_scores:  
Quality: 48.00 Length: 48  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q92311

Align seg 1/1 to: Q92311 from: 1 to: 1050

2815 GAGCTGTGGCCAACTGGGAAGACCCCAAGAGAGAAACATTTTAAATTATG 2864

```
|||||
912 GluLeuValAlaLysLeuGluAspProArgGluLysHisPheAsnLeuCy 928
2865 TCTCGAGGAAAGGACTGGTTACCAGGGCAGCCAGTCTCTCGAAAAACCTTT 2914
|||||
928 sLeuGluGluArgAspTrpLeuProGlyGlnProValLeuGluAsnLeuS 945
|||||
2915 CCAGAGCATACAGCTTAGCAAAAAGACAGTGTGTGTGATGACA 2958
|||||
945 erGlnSerIleGlnLeuSerLysLysThrValPheValMetThr 959
|||||
```

seq\_name: sp\_rodent:Q91X17

seq\_documentation\_block:

ID Q91X17 PRELIMINARY; PRT; 1032 AA.

AC Q91X17;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 8.

GN TLR8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE-SPLEEN;  
RA Heil F.J., Lipford G.B., Wagner H., Bauer S.M.;  
RT "Molecular cloning of murine Toll-Like Receptor 8";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035890; AAK62677.1; -;  
KW Receptor.  
SQ SEQUENCE 1032 AA; 119338 MW; B2812191D99EE42D CRC64;

alignment\_scores:

Quality: 17.00 Length: 17  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q91X17

Align seg 1/1 to: Q91X17 from: 1 to: 1032

844 CTTGACCTAAGTGGAAATGCCCTCGTTGTATTAATGCCCATTTCTCTTG 893  
|||||

246 LeuAspLeuSerGlyAsnCysProArgCysTyraAsnAlaPropheProCy 262  
|||||

894 T 894

262 s 262

seq\_name: sp\_human:Q9H5G9

seq\_documentation\_block:

ID Q9H5G9 PRELIMINARY; PRT; 363 AA.

AC Q9H5G9;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CDNA: FLJ23447 FIS, CLONE HSI03346.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE-SMALL INTESTINE;  
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,

RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
RA Isogai T., Sugano S.;  
RT "NEDO human cDNA sequencing project.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK027100; BAB15657.1; -;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR003591; LRR\_tyr.  
DR Pfam: PF00560; LRR; 8.  
DR PRINTS: PR00109; LEURICHRPT.  
DR SMART: SM00370; LRR; 2.  
DR SMART: SM00369; LRR\_TYP; 8.  
SQ SEQUENCE 363 AA; 39926 MW; DD6999A339228C76 CRC64;

alignment\_scores:  
Quality: 12.00 Length: 12  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9H5G9 ..

Align seg 1/1 to: Q9H5G9 from: 1 to: 363

2194 TTGGACCTCAGCCACACCACTGACCACTGTCCCT 2229  
|||||  
122 LeuAspLeuSerHisAsnGlnLeuThrValPro 133

seq\_name: sp\_plant:Q9SN91

seq\_documentation\_block:

ID Q9SN91 PRELIMINARY; PRT; 1232 AA.

AC Q9SN91;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE LEUCINE RICH REPEAT-LIKE PROTEIN.  
GN FIC12.60 OR AT4G20140.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Terryn N., Ardiles W., Buysshaert C., Dasseville R.,  
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
RA Villaroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,  
RA Mayer K.F.X., Lemcke K., Schueller C.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1-305 FROM N.A.  
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Terryn N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,  
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,  
RA Gielen J., Van Montagu M., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL: AL022224; CA18239.1; -;  
DR EMBL: AL161552; CAB79014.1; -;  
DR HSSP: P08631; IAD5.

DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00560; LRR; 30.  
DR Pfam: PF00089; pkinase; 1.  
DR PRINTS: PR00019; LEURICHRPT.  
DR PRINTS: PR00109; TYRKINASE.  
DR SMART: SM00370; LRR; 30.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 1232 AA; 135534 MW; EAD6322CE8967726 CRC64;

alignment\_scores:  
Quality: 12.00 Length: 12  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9SN91 ..

Align seg 1/1 to: Q9SN91 from: 1 to: 1232

2185 CTGGAACCTTTGGACCTCAGCCACACCACTGACC 2220  
|||||  
794 LeuGluThrLeuAspLeuSerHisAsnGlnLeuThr 805

seq\_name: sp\_human:Q9NYC2

seq\_documentation\_block:

ID Q9NYC2 PRELIMINARY; PRT; 975 AA.  
AC Q9NYC2;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 9 FORM B.  
GN TLR9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20477806; Pubmed=11022119;  
RA Du X., Poltorak A., Wei Y., Beutler B.;  
RT "Three novel mammalian toll-like receptors: gene structure,  
RT expression, and evolution.";  
RL Eur. Cytokine Netw. 11:362-371(2000).  
DR EMBL: AF259263; AAF72190.1; -;  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR003591; LRR\_tyr.  
DR InterPro: IPR000157; TIR.  
DR Pfam: PF00560; LRR; 17.  
DR Pfam: PF01582; TIR; 1.  
DR PRINTS: PR00019; LEURICHRPT.  
DR SMART: SM00370; LRR; 4.  
DR SMART: SM00369; LRR\_TYP; 2.  
DR SMART: SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 975 AA; 109628 MW; 61D05163587A75F CRC64;

alignment\_scores:  
Quality: 11.00 Length: 11  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

```
US-09-202-054-2 x Q9NYC2
Align seg 1/1 to: Q9NYC2 from: 1 to: 975
2860 TTATGCTCTCGAGAAAGGACTGGTTACCAGG 2892
|||||
849 LeuCysLeuGluGluArgAspTrpLeuProGly 859
seq_name: sp_human:Q9NYC3
seq_documentation_block:
ID Q9NYC3 PRELIMINARY; PRT; 1032 AA.
AC Q9NYC3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TOLL-LIKE RECEPTOR 9 FORM A.
GN TLR9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PLACENTA;
RA Chuang T.H., Ulevitch R.J.;
MEDLINE=20477807; PubMed=11022120;
RT "Cloning and characterization of a sub-family of human toll-like
RL receptors: hTLR7, hTLR8 and hTLR9."
DR EMBL: AF245704; AAF78037.1; -.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00560; LRR; 17.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR; 4.
DR SMART: SM00369; LRR_Typ; 2.
DR SMART: SM00255; TIR; 1.
KW Receptor.
SQ SEQUENCE 1032 AA; 115906 MW; 2B053A8595FDC9FF CRC64;

alignment_scores:
Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Q9NR96
Align seg 1/1 to: Q9NR96 from: 1 to: 1032
2860 TTATGCTCTCGAGAAAGGACTGGTTACCAGG 2892
|||||
906 LeuCysLeuGluGluArgAspTrpLeuProGly 916
seq_name: sp_fungi:Q08817
seq_documentation_block:
ID Q08817 PRELIMINARY; PRT; 791 AA.
AC Q08817;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CHROMOSOME XV READING FRAME ORF YOR353C.
GN YOR353C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z75261; CAA99682.1; -.
DR SGD: S0005880; YOR353C.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR Pfam: PF00560; LRR; 4.
DR SMART: SM00370; LRR; 2.
SQ SEQUENCE 791 AA; 87325 MW; B0EA559AA4F66199 CRC64;

alignment_scores:
Quality: 11.00 Length: 11
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Q9NYC3
Align seg 1/1 to: Q9NYC3 from: 1 to: 1032
2860 TTATGCTCTCGAGAAAGGACTGGTTACCAGG 2892
|||||
906 LeuCysLeuGluGluArgAspTrpLeuProGly 916
seq_name: sp_human:Q9NR96
seq_documentation_block:
ID Q9NR96 PRELIMINARY; PRT; 1032 AA.
AC Q9NR96;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TOLL-LIKE RECEPTOR 9
```

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q08817 ..

Align seg 1/1 to: Q08817 from: 1 to: 791

706 GTGCTCTCCCTGAAGATAACATGTCACA 735

|||||  
141 valLeuSerLeuLysAspAsnValThr 150

seq\_name: sp\_human:Q9NR97

seq\_documentation\_block:

ID Q9NR97 PRELIMINARY; PRT; 1041 AA.

AC Q9NR97;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE TOLL-LIKE RECEPTOR 8.

GN TLR8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=20477807; PubMed=11022120;

RA Chuang T.H., Ulevitch R.J.;

RT "Cloning and characterization of a sub-family of human toll-like

receptors: hTLR7, hTLR8 and hTLR9.";

RL Eur. Cytokine Netw. 11:372-378(2000).

DR EMBL; AF245703; AAF78036.1; -.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR InterPro; IPR000157; TIR.

DR Pfam; PF00560; LRR; 16.

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01582; TIR; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 3.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_TYP; 3.

DR SMART; SM00255; TIR; 1.

KW Receptor.

SQ SEQUENCE 1041 AA; 119827 MW; 39A38B60629291C8 CRC64;

alignment\_scores:  
Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9NR97 ..

Align seg 1/1 to: Q9NR97 from: 1 to: 1041

844 CTTGACCTAAGTGAATGCCCTCGTTGT 873

|||||  
251 LeuAspLeuSerGlyAsnCysProArgCys 260

seq\_name: sp\_human:Q9NYG9

seq\_documentation\_block:

ID Q9NYG9 PRELIMINARY; PRT; 1059 AA.

AC Q9NYG9;

DT 01-OCT-2000-(TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE TOLL-LIKE RECEPTOR 8.  
GN TLR8.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=20477806; PubMed=11022119;

RA Du X., Poltorak A., Wei Y., Beutler B.;

RT "Three novel mammalian toll-like receptors: gene structure,

expression, and evolution.";

RL Eur. Cytokine Netw. 11:362-371(2000).

DR EMBL; AF246971; AAF64061.1; -.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR InterPro; IPR000157; TIR.

DR Pfam; PF00560; LRR; 16.

DR Pfam; PF01463; LRRCT; 1.

DR Pfam; PF01582; TIR; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 4.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_TYP; 2.

DR SMART; SM00255; TIR; 1.

KW Receptor.

SQ SEQUENCE 1059 AA; 121751 MW; 2025AEB6DBB7C4AF CRC64;

alignment\_scores:

Quality: 10.00 Length: 10  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9NYG9 ..

Align seg 1/1 to: Q9NYG9 from: 1 to: 1059

844 CTTGACCTAAGTGAATGCCCTCGTTGT 873

|||||  
269 LeuAspLeuSerGlyAsnCysProArgCys 278

seq\_name: sp\_rodent:Q63156

seq\_documentation\_block:

ID Q63156 PRELIMINARY; PRT; 96 AA.

AC Q63156;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE DECORIN (FRAGMENT).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;

RA Moats-Staats B.M., Stiles A.D., Xu L.;

RT "Expression of decorin RNA in rat lung undergoing chronic lung

injury.";

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; L75825; AAA85371.1; -.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003592; LRR\_out.

DR Pfam; PF00560; LRR; 3.

DR SMART; SM00370; LRR; 2.  
 FT NON\_TER 1 1  
 FT NON\_TER 96 96  
 SQ SEQUENCE 96 AA; 10398 MW; C6D19F1750B050D0 CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-202-054-2 x Q63156 ..

Align seg 1/1 to: Q63156 from: 1 to: 96

508 ATACCGAGGGCTCCCGCTAGCTTA 534  
 |||||  
 12 IleProGlnGlyLeuProSerLeu 20

seq\_name: sp\_human:Q9P0Z1

seq\_documentation\_block:  
 ID Q9P0Z1 PRELIMINARY; PRT; 250 AA.  
 AC Q9P0Z1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE DECORIN B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cs-Szabo G., Glant T.T.;  
 RT "Alternative splicing of human decorin.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF138301; AAF61437.1; ..  
 DR InterPro: IPR001611; LRR\_Nterm.  
 DR InterPro: IPR003572; LRR\_out.  
 DR InterPro: IPR003592; LRR\_Nterm.  
 DR Pfam: PF00560; LRR; 3.  
 DR Pfam: PF01462; LRRNT; 1.  
 DR SMART; SM00370; LRR; 2.  
 DR SMART; SM00013; LRRNT; 1.  
 SQ SEQUENCE 250 AA; 27353 MW; 5AA599BE479F68D9 CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-202-054-2 x Q9P0Z1 ..

Align seg 1/1 to: Q9P0Z1 from: 1 to: 250

508 ATACCGAGGGCTCCCGCTAGCTTA 534  
 |||||  
 106 IleProGlnGlyLeuProSerLeu 114

seq\_name: sp\_plant:Q94L68

seq\_documentation\_block:  
 ID Q94L68 PRELIMINARY; PRT; 252 AA.  
 AC Q94L68;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE POLYGALACTURONASE INHIBITOR PROTEIN (FRAGMENT).  
 GN PGIP.

OS Potentilla anserina.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Rosales; Rosaceae; Rosoideae; Potentilla.  
 OX NCBI\_TaxID=57926;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PO\_AN.7;  
 RA Potter D., Oh S.-H., Gao F., Baggett S.;  
 RT "Phylogenetic relationships among putative genes encoding  
 RT polygalacturonase inhibitor proteins (PGIPs) in Rosaceae.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF196917; AAK43431.1; ..  
 FT NON\_TER 1 1  
 FT NON\_TER 252 252  
 SQ SEQUENCE 252 AA; 28108 MW; 3CA7578D862DDCC6 CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-202-054-2 x Q94L68 ..

Align seg 1/1 to: Q94L68 from: 1 to: 252

2194 TTGGACCTCAGCCACCAACTGACC 2220  
 |||||  
 150 LeuAspLeuSerHisAsnGlnLeuThr 158

seq\_name: sp\_plant:Q94L68

seq\_documentation\_block:  
 ID Q94L68 PRELIMINARY; PRT; 252 AA.  
 AC Q94L68;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE POLYGALACTURONASE INHIBITOR PROTEIN (FRAGMENT).  
 GN PGIP.  
 OS Potentilla anserina.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Rosales; Rosaceae; Rosoideae; Potentilla.  
 OX NCBI\_TaxID=57926;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PO\_AN.7;  
 RA Potter D., Oh S.-H., Gao F., Baggett S.;  
 RT "Phylogenetic relationships among putative genes encoding  
 RT polygalacturonase inhibitor proteins (PGIPs) in Rosaceae.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF196917; AAK43431.1; ..  
 FT NON\_TER 1 1  
 FT NON\_TER 252 252  
 SQ SEQUENCE 252 AA; 28078 MW; 39F5C458D80DA380 CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-202-054-2 x Q94L68 ..

Align seg 1/1 to: Q94L68 from: 1 to: 252

2194 TTGGACCTCAGCCACCAACTGACC 2220  
 |||||  
 150 LeuAspLeuSerHisAsnGlnLeuThr 158



seq\_name: sp\_invertebrate:Q9XYQ8

```
seq_documentation_block:
ID Q9XYQ8 PRELIMINARY; PRT; 289 AA.
AC Q9XYQ8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BONE MORPHOGENETIC PROTEIN BMP2/4 (FRAGMENT).
GN BMP2/4.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RA Angerer L.M., Oleksyn D.W., Logan C.Y., McClay D.R., Dale L.,
RA Angerer R.C.;
RT "A BMP pathway regulates cell fate allocation along the sea urchin
animal-vegetal embryonic axis.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF119713; AAD28039.1; -.
DR HSSP; PI2643; 3BMP.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR002405; Inhibin.alpha.
DR InterPro; IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta.1.
DR PRINTS; PR00438; GFCYSNOT.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGF-beta; 1.
DR SMART; SM00204; TGF-beta; 1.
DR PROSITE; PS00250; TGF-BETA; 1.
KW Glycoprotein.
FT NON_TER 1
FT NON_TER 289
SQ SEQUENCE 289 AA; 33599 MW; 99E175C7DBC3C58B CRC64;
```

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9XYQ8 ..

Align seg 1/1 to: Q9XYQ8 from: 1 to: 289

```
3071 AGAAGTCCCAAGTTCCTCCAGCTCCGGA 3097
|||||
163 ArgSerProSerSerSerSerSerGly 171
```

seq\_name: sp\_plant:Q9ZSK1

```
seq_documentation_block:
ID Q9ZSK1 PRELIMINARY; PRT; 348 AA.
AC Q9ZSK1;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GAMMA-TOCOPHEROL METHYLTRANSFERASE.
GN G-TWT.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shintani D.K., DellaPenna D.;
```

```
RT "Elevating Vitamin E content of plants through metabolic
engineering.";
RL Science 0:0-0(1998).
DR EMBL; AF104220; AAD02882.1; -.
DR InterPro; IPR001601; Meth-transf.
DR InterPro; IPR000051; SAM_bind.
KW Transferase; Methyltransferase.
SQ SEQUENCE 348 AA; 38091 MW; 44DB18A722E0725F CRC64;
```

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9ZSK1 ..

Align seg 1/1 to: Q9ZSK1 from: 1 to: 348

```
3068 TTCAGAGTCCCAAGTTCCTCCAGCTCC 3094
|||||
30 PreArgSerProSerSerSerSerSer 38
```

seq\_name: sp\_plant:Q9XIP9

seq\_documentation\_block:

ID Q9XIP9 PRELIMINARY; PRT; 348 AA.

AC Q9XIP9;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE GAMMA-TOCOPHEROL METHYLTRANSFERASE (AT1G64970/F13011\_27).

GN F13011\_27

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,

RA Altafi H., Araujo R., Huijar L., Rowley D., Buehler E., Dunn P.,

RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,

RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,

RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,

RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,

RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,

RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,

RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,

RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,

RA Ecker J.R.;

RT "Arabidopsis cDNA clones";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006193; AAD38271.2; -.

DR EMBL; AY049258; AAK83600.1; -.

DR InterPro; IPR001601; Meth-transf.

DR InterPro; IPR000051; SAM\_bind.

KW Transferase; Methyltransferase.

SQ SEQUENCE 348 AA; 38075 MW; E9290758C2E83B73 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x Q9XIP9 ..

Align seg 1/1 to: Q9XIP9 from: 1 to: 348

3068 TTCAGAAAGTCCAAAGTCTCCAGCTCC 3094

|||||

30 PheArgSerProSerSerSerSer 38

seq\_name: sp\_plant:Q9LMR3

## seq\_documentation\_block:

ID Q9LMR3 PRELIMINARY; PRT; 358 AA.

AC Q9LMR3;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE F7H2.5 PROTEIN (AROGENATE DEHYDROGENASE ISOFORM 2) (EC 1.3.1.43).

GN F7H2.5.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV, COLUMBIA;

RA Liu S.X., Sakano H., Yu G., Etgu P., Lee J., Lenz C., Pham P.,

RA Toriumi M., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,

RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,

RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,

RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,

RA Ecker J.R., Federspiel N.A., Theologis A.;

RT "The sequence of BAC F7H2 from Arabidopsis thaliana chromosome 1.";

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Rippert P., Matringe M.;

RT "Arabidopsis thaliana arogenate dehydrogenase mRNA isoform 2.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC034256; AAL2141.1; -.

DR EMBL; AF434682; AAL30406.1; -.

KW Oxidoreductase.

SQ SEQUENCE 358 AA; 40633 MW; DCEFA72C62D1AF78 CRC64;

## alignment\_scores:

Quality: 9.00 Length: 9

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2/rev x Q9LMR3 ..

Align seg 1/1 to: Q9LMR3 from: 1 to: 358

2317 CTGTGAGAAATACCTCGTCAGACTCC 2291

|||||

268 LeuValGluAsnThrSerSerAspSer 276

seq\_name: sp\_mammal:Q9TTE2

## seq\_documentation\_block:

ID Q9TTE2 PRELIMINARY; PRT; 360 AA.

AC Q9TTE2;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DECORIN

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI\_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=MYOMETRIUM;

RX MEDLINE=20113292; PubMed=10644528;

RA Wu W.X., Zhang Q., Unno N., Derks J.B., Nathanielsz P.W.;

RT "Characterization of decorin mRNA in pregnant intrauterine tissues of

RT the ewe and regulation by steroids.";

RL Am. J. Physiol. 278:C199-C206(2000).

DR EMBL; AF125041; AAF00585.1; -.

DR HSSP; P09661; IAGN.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000372; LRR\_Nterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR Pfam; PF00560; LRR; 9.

DR Pfam; PF01462; LRRNT; 1.

DR SMART; SM00370; LRR; 2.

DR SMART; SM00013; LRRNT; 1.

DR SMART; SM00369; LRR\_TYP; 2.

SQ SEQUENCE 360 AA; 39972 MW; 0095D0DFDAB88624 CRC64;

## alignment\_scores:

Quality: 9.00

Ratio: 1.000

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x Q9TTE2 ..

Align seg 1/1 to: Q9TTE2 from: 1 to: 360

508 ATACCGCAGGCGCTCCGCTAGCTTA 534

|||||

216 IleProGlnGlyLeuProSerLeu 224

seq\_name: sp\_human:Q9ULQ7

## seq\_documentation\_block:

ID Q9ULQ7 PRELIMINARY; PRT; 437 AA.

AC Q9ULQ7;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE KIAA1163 PROTEIN (FRAGMENT).

GN KIAA1163.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=20039618; PubMed=10574461;

RA Hirosawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;

RT "Characterization of cDNA clones selected by the Genemark analysis

RT from size-fractionated cDNA libraries from human brain.";

RL DNA Res. 6:329-336(1999).

DR EMBL; AB032989; BAA86477.1; -.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR Pfam; PF00047; Ig; 1.

DR Pfam; PF00560; LRR; 5.

DR Pfam; PF01463; LRRCT; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00409; IG; 1.

DR SMART; SM00370; LRR; 3.

DR SMART; SM00082; LRRCT; 1.

DR SMART: SM00369; LRR\_TYP: 2.  
 FT NON\_TER 1  
 SQ SEQUENCE 437 AA; 49206 MW; 23672C9873E81757 CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9ULQ7 ..

Align seg 1/1 to: Q9ULQ7 from: 1 to: 437

2329 TTGCGATATCTGATCTCAGCTCAAAAT 2355

|||||  
 56 LeuArgTyrLeuAspLeuSerSerAsn 64

seq\_name: sp\_human:Q9ULM6

seq\_documentation\_block:

ID Q9ULM6 PRELIMINARY; PRT; 575 AA.

AC Q9ULM6

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)

DE KIAA1194

GN KIAA1194

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=20039619; PubMed=10574462;

RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XV.

RT The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro.;"

RL DNA Res. 6:337-345(1999).

DR EMBL; AB033020; BAA86508.1; ..

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR Pfam; PF00560; LRR; 3.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 1.

DR SMART; SM00369; LRR\_TYP; 2.

FT NON\_TER 1

SQ SEQUENCE 575 AA; 65384 MW; 2B2748A8A4852C68 CRC64;

alignment\_scores:

Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9ULM6 ..

Align seg 1/1 to: Q9ULM6 from: 1 to: 575

2335 TATCTGATCTCAGCTCAATAAATC 2361

|||||

96 TyrLeuAspLeuSerSerAsnLysIle 104

seq\_name: sp\_human:Q9HDC7

seq\_documentation\_block:

ID Q9HDC7 PRELIMINARY; PRT; 626 AA.

AC Q9HDC7;

DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE PLATELET GLYCOPROTEIN IB ALPHA.  
 GN GPIB ALPHA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matsubara Y., Murata M., Moriki T., Yokoyama K., Watanabe N.,  
 RA Nakajima H., Handa M., Ikeda Y.;  
 RT "A new polymorphism, 70Leu/Phe, within the leucine-rich repeat  
 RT sequence of platelet glycoprotein IB-alpha.;"  
 RL Blood 0:0-0(2000).  
 DR EMBL; AB038516; BAB12038.1; ..  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003592; LRR\_out.  
 DR InterPro; IPR003591; LRR\_typ.  
 DR InterPro; IPR002965; P\_rich\_extensn.  
 DR Pfam; PF00560; LRR; 6.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR SMART; SM00370; LRR; 3.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 5.  
 SQ SEQUENCE 626 AA; 68989 MW; 4DB14119B742D222 CRC64;

alignment\_scores:

Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9HDC7 ..

Align seg 1/1 to: Q9HDC7 from: 1 to: 626

2191 ACTTTGGACCTCAGCCACCAACCTG 2217

|||||

97 ThrLeuAspLeuSerHisAsnGlnLeu 105

seq\_name: sp\_rodent:Q921U9

seq\_documentation\_block:

ID Q921U9 PRELIMINARY; PRT; 626 AA.

AC Q921U9;

DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE UNKNOWN (PROTEIN FOR IMAGE:3498778) (FRAGMENT).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC010598; AAI10598.1; ..

FT NON\_TER 1

SQ SEQUENCE 626 AA; 69123 MW; 5AF3570E270A2DFF CRC64;

alignment\_scores:

Quality: 9.00 Length: 9

Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x Q921U9 ..

Align seg 1/1 to: Q921U9 from: 1 to: 626

2329 TTGGATATCTGGATCTCAAGTCAAAAT 2355

|||||

246 LeuArgTyLeuAspLeuSerSerAsn 254

seq\_name: sp\_human:Q9UK78

## seq\_documentation\_block:

ID Q9UK78 PRELIMINARY; PRT; 752 AA.

AC Q9UK78

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE HUMAN TOLL-LIKE RECEPTOR 4 (FRAGMENT).

GN TLR4

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Poltorak A., Smirnova I., Beutler B.;

RT "Genetic variation at the TLR4 locus.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF177766; AAF07823.1; -.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR000157; LRR\_typ.

DR Pfam; PF00560; LRR; 7.

DR Pfam; PF01463; LRRCT; 1.

DR PRINTS; PR001582; TIR; 1.

DR SMART; SM00370; LRR; 1.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_TYP; 2.

DR SMART; SM00255; TIR; 1.

KW Receptor.

FT NON\_TER

SQ SEQUENCE 752 AA; 85715 MW; 3275C96C06E1A1A2C CRC64;

## alignment\_scores:

Quality: 9.00

Ratio: 1.000

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x Q9UK78 ..

Align seg 1/1 to: Q9UK78 from: 1 to: 752

2338 CTGGATCTCAGCTCAATAAATCCAG 2364

|||||

93 LeuAspLeuSerSerAsnLysIleGln 101

seq\_name: sp\_human:Q9UM57

## seq\_documentation\_block:

ID Q9UM57 PRELIMINARY; PRT; 799 AA.

AC Q9UM57

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE TOLL-LIKE RECEPTOR 4.

GN TLR4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98118556; PubMed=9435236;

RA Rock F.L., Hardman G., Timans J.C., Kastelein R.A., Bazan J.F.;

RT "A family of human receptors structurally related to Drosophila

TLR4."

RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).

DR EMBL; U88880; AAC34135.1; -.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR000157; LRR\_typ.

DR Pfam; PF00560; LRR; 9.

DR Pfam; PF01463; LRRCT; 1.

DR PRINTS; PR001582; TIR; 1.

DR SMART; SM00370; LRR; 2.

DR SMART; SM00082; LRRCT; 1.

DR SMART; SM00369; LRR\_TYP; 2.

DR SMART; SM00255; TIR; 1.

KW Receptor.

SQ SEQUENCE 799 AA; 91295 MW; 82F70995E7F2AF9D CRC64;

## alignment\_scores:

Quality: 9.00

Ratio: 1.000

Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-202-054-2 x Q9UM57 ..

Align seg 1/1 to: Q9UM57 from: 1 to: 799

2338 CTGGATCTCAGCTCAATAAATCCAG 2364

|||||

140 LeuAspLeuSerSerAsnLysIleGln 148

seq\_name: sp\_mammal:Q9TSP2

## seq\_documentation\_block:

ID Q9TSP2 PRELIMINARY; PRT; 826 AA.

AC Q9TSP2;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE TOLL-LIKE RECEPTOR 4.

GN TLR4.

OS Papio anubis (Olive baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

OC Cercopitheidae; Papio.

OX NCBI\_TaxID=9555;

RN [1]

RP SEQUENCE FROM N.A.

RA Smirnova I., Poltorak A., Chan E.K.L., Alejos E., McBride C.,

RA Beutler B.;

RT "Genetic variation at the TLR4 locus.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF180964; AAF07059.1; -.

DR EMBL; AF180962; AAF07059.1; JOINED.

DR EMBL; AF180963; AAF07059.1; JOINED.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR000483; LRR\_Cterm.

DR InterPro; IPR003592; LRR\_out.

DR InterPro; IPR003591; LRR\_typ.

DR InterPro; IPR000157; TIR.

DR Pfam: PF00560; LRR; 9.  
DR Pfam: PF01463; LRRCT; 1.  
DR Pfam: PF01582; TIR; 1.  
DR PRINTS: PR00019; LEURICHRPT.  
DR SMART: SM00370; LRR; 1.  
DR SMART: SM00082; LRRCT; 1.  
DR SMART: SM00369; LRR\_TYP; 2.  
DR SMART: SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 826 AA; 94678 MW; 422777318E5F1769 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9TSP2 ..

Align seg 1/1 to: Q9TSP2 from: 1 to: 826

2338 CTGGATCTCAGCTCAATAAAATCCAG 2364

|||||

180 LeuAspLeuSerSerAsnLysIleGln 188

seq\_name: sp\_human:000206

seq\_documentation\_block:

ID O00206 PRELIMINARY; PRT; 839 AA.

AC O00206;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE TOLL PROTEIN HOMOLOG (TOLL-LIKE RECEPTOR 4).

GN TLR4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97379437; PubMed=9237759;

RA Medzhitov R., Preston-Hurlburt P., Janeway C.A.;

RT "A human homologue of the Drosophila Toll protein signals activation

RT of adaptive immunity.";

RL Nature 388:394-397(1997).

RN [2]

RP SEQUENCE FROM N.A.

RA Poltorak A., Smirnova I., Chan E.K.L., Beutler B.;

RT "Genetic variation at the TLR4 locus.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Arbour N.C., Lorenz E., Schutte B.C., Zabner J., Kline J.N., Jones M.,

RA Fries K., Watt J.L., Schwartz D.A.;

RT "A Genetic Basis for a Blunted Response to Endotoxin in Humans.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: U93091; AAC80227.1;

DR EMBL: AF177765; AAF05316.1;

DR EMBL: AF172171; AAF89753.1;

DR EMBL: AF172169; AAF89753.1; JOINED.

DR EMBL: AF172170; AAF89753.1; JOINED.

DR InterPro: IPR001611; LRR.

DR InterPro: IPR000483; LRR\_Cterm.

DR InterPro: IPR003592; LRR\_Out.

DR InterPro: IPR003591; LRR\_Type.

DR InterPro: IPR000157; TIR.

DR Pfam: PF00560; LRR; 9.

DR Pfam: PF01463; LRRCT; 1.

DR Pfam: PF01582; TIR; 1.

DR PRINTS: PR00019; LEURICHRPT.

DR SMART: SM00370; LRR; 2.

DR SMART: SM00082; LRRCT; 1.  
DR SMART: SM00369; LRR\_TYP; 2.  
DR SMART: SM00255; TIR; 1.  
KW Receptor.  
SQ SEQUENCE 839 AA; 95679 MW; 92C48F55821133E8 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x O00206 ..

Align seg 1/1 to: O00206 from: 1 to: 839.

2338 CTGGATCTCAGCTCAATAAAATCCAG 2364

|||||

180 LeuAspLeuSerSerAsnLysIleGln 188

seq\_name: sp\_mammal:Q9TTN0

seq\_documentation\_block:

ID Q9TTN0 PRELIMINARY; PRT; 839 AA.

AC Q9TTN0;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE TOLL-LIKE RECEPTOR 4.

GN TLR4.

OS Pan paniscus (Pygmy chimpanzee) (Bonobo).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

OX NCBI\_TaxID=9597;

RN [1]

RP SEQUENCE FROM N.A.

RA Smirnova I., Poltorak A., Chan E.K.L., Alejos E., McBride C.,

RA Beutler B.;

RT "Genetic variation at the TLR4 locus.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF179220; AAF05320.1;

DR EMBL: AF179218; AAF05320.1; JOINED.

DR EMBL: AF179219; AAF05320.1; JOINED.

DR InterPro: IPR001611; LRR.

DR InterPro: IPR000483; LRR\_Cterm.

DR InterPro: IPR003592; LRR\_Out.

DR InterPro: IPR003591; LRR\_Type.

DR Pfam: PF00560; LRR; 9.

DR Pfam: PF01463; LRRCT; 1.

DR Pfam: PF01582; TIR; 1.

DR PRINTS: PR00019; LEURICHRPT.

DR SMART: SM00370; LRR; 2.

DR SMART: SM00082; LRRCT; 1.

DR SMART: SM00369; LRR\_TYP; 2.

DR SMART: SM00255; TIR; 1.

KW Receptor.

SQ SEQUENCE 839 AA; 95637 MW; 3B328C5682127D37 CRC64;

alignment\_scores:

Quality: 9.00 Length: 9

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9TTN0 ..

Align seg 1/1 to: Q9TTN0 from: 1 to: 839

2338 CTGGATCTCAGCTCAATAAAATCCAG 2364

|||||

```
180 LeuAspLeuSerSerAsnIleGln 188
seq_name: sp_plant:Q9LNX8
seq_documentation_block:
ID Q9LNX8 PRELIMINARY; PRT; 945 AA.
AC Q9LNX8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F22G5.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chioi J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F22G5 from chromosome
I.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC022464; AAF9546.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00370; LRR; 2.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS50107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 945 AA; 106684 MW; C6DC79C69B69727E CRC64;

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Q9LNX8 ..
Align seg 1/1 to: Q9LNX8 from: 1 to: 945
1030 CTCAGGAAGTGGATCTGTCCTCCCAAC 1056
|||||
461 LeuGlnGluLeuAspLeuSerGlnAsn 469

seq_name: sp_invertebrate:Q9V701
seq_documentation_block:
ID Q9V701 PRELIMINARY; PRT; 953 AA.
AC Q9V701;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG8561 PROTEIN.
RN [1]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dutilleul K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fouts D.E., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Klotzel C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paciorek J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003815; AAF38265.1; -.
DR FlyBase; FBgn0033920; CG8561.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 22.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 9.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00369; LRR_TYP; 3.
SQ SEQUENCE 953 AA; 108032 MW; 16D4C22AD854756B CRC64;

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Q9V701 ..
Align seg 1/1 to: Q9V701 from: 1 to: 953
2185 CTGGAAGACTTTGGACCTCAGCCCAAC 2211
```

```

|||||
626 LeuGluThrLeuAspLeuSerHisAsn 634
seq_name: sp_plant:Q9LKZ4

seq_documentation_block:
ID Q9LKZ4 PRELIMINARY; PRT; 1012 AA.
AC Q9LKZ4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE RECEPTOR-LIKE PROTEIN KINASE 3.
GN RLK3.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21363855; PubMed=11470843;
RY Yamamoto E., Knap H.T.;
RT "Soybean receptor-like protein kinase genes: paralogous divergence of
a gene family.";
RL Mol. Biol. Evol. 18:1522-1531(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF244890; AAF91324.1; -.
DR HSSP; P12931; 1FWK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF00560; LRR; 21.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINRPT.
DR PRINTS; PR00109; TYRKINRPT.
DR SMART; SM00370; LRR; 18.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1012 AA; 110322 MW; FD555FB57F99815D CRC64;

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Q9LKZ4 ..
Align seg 1/1 to: Q9LKZ4 from: 1 to: 1012

1270 CITCAAAATCTTGAAGTCTTGTGATCTT 1296
|||||
138 LeuGlnAsnLeuGluValLeuAspLeu 146
seq_name: sp_rodent:Q9JJ28

seq_documentation_block:
ID Q9JJ28 PRELIMINARY; PRT; 1271 AA.
AC Q9JJ28;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE FLIIH PROTEIN.
GN FLIIH.
OS Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C;
RY Campbell H.D., Fountain S., Young I.G., Weitz S., Lichter P.,
RT Hoheisel J.D.;
RT "FlIIH, the murine homologue of the Drosophila melanogaster flightless
I gene: nucleotide sequence, chromosomal mapping and overlap with
LigIh.";
RL DNA Seq. 11:29-40(2000).
DR EMBL; AF142329; AAF78453.1; -.
DR HSSP; P02640; 2VIL.
DR MGD; MGI:1342286; FlIIh.
DR InterPro; IPR001974; Gelsolin.
DR InterPro; IPR003015; HLH_Myc.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR Pfam; PF00626; Gelsolin; 5.
DR Pfam; PF00560; LRR; 12.
DR PRINTS; PR00597; GELSOLIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00262; GEL; 6.
DR SMART; SM00370; LRR; 7.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 1271 AA; 144802 MW; A9642B10FEBF8769 CRC64;

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x Q9JJ28 ..
Align seg 1/1 to: Q9JJ28 from: 1 to: 1271

2194 TTGGACCTCAGCCACCAACCACTGACC 2220
|||||
109 LeuAspLeuSerHisAsnGlnLeuThr 117
seq_name: sp_invertebrate:Q9VGII

seq_documentation_block:
ID Q9VGII PRELIMINARY; PRT; 1354 AA.
AC Q9VGII;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CGI7360 PROTEIN.
GN CGI7360.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RY MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE036933; AAF34698.1; -.  
 DR FlyBase; FBgn0037949; CG17360.  
 DR InterPro; IPR001849; PH.  
 DR Pfam; PF00169; PH; 1.  
 DR SMART; SM00233; PH; 1.  
 DR PROSITE; PS50003; PH\_DOMAIN; 1.  
 SQ SEQUENCE 1354 AA; 149694 MW; C5D275B5CB9E2695 CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x Q9VGII ..

Align seg 1/1 to: Q9VGII from: 1 to: 1354

2267 AGATCTTGAGGCTCTGGAACAGTTG 2241

|||||  
 1302 ArgPheLeuArgLeuLeuGluGlnLeu 1310

seq\_name: sp\_plant:Q9C6R1

seq\_documentation\_block:

ID Q9C6R1 PRELIMINARY; PRT: 1784 AA.

AC Q9C6R1;

DT 01-JUN-2001 (TRENBLrel. 17, Created)

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE HYPOTHETICAL 201.8 KDA PROTEIN.

GN T18124.10.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=21016719; PubMed=11130712;

RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,

RA White O., Alonso J., Alcafi H., Araujo R., Bowman C.L., Brooks S.Y.,

RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,

RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.-Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langan-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*  
 thaliana".  
 RL Nature 408:816-820(2000).  
 DR EMBL; AC079131; AAG50756.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003591; LRR\_type.  
 DR Pfam; PF00560; LRR; 44.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00369; LRR\_TYP; 20.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1784 AA; 201803 MW; 98AEB6FFD6AC8F9D CRC64;

alignment\_scores:  
 Quality: 9.00 Length: 9  
 Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9C6R1 ..

Align seg 1/1 to: Q9C6R1 from: 1 to: 1784

1030 CTCAGGACTGCTGTCCCAAAAC 1056

|||||

226 LeuGlnGluLeuAspLeuSerGlnAsn 234

seq\_name: sp\_invertebrate:Q19761

seq\_documentation\_block:

ID Q19761 PRELIMINARY; PRT: 2287 AA.

AC Q19761; Q22487;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)

DE F23D12.2 PROTEIN.

GN F23D12.2

OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; *Caenorhabditis*.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Barlow K.

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,

RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

RA Watson A., Welnscock L., Wilkinson-Sproat J., Wohlman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*.";



RL Nature 368:32-38(1994).  
DR EMBL; Z71186; CAA94917.1; -.  
DR EMBL; Z70687; CAA94917.1; JOINED.  
DR EMBL; Z70687; CAA94618.1; -.  
DR EMBL; Z71186; CAA94618.1; JOINED.  
DR InterPro; IPR001810; F-box.  
DR Pfam; PF00646; F-box; 1.  
DR SMART; SM00256; FBOX; 1.  
DR PROSITE; PS0181; FBOX; 1.  
SQ SEQUENCE 2287 AA; 255336 MW; 3496918CC952EF56 CRC64;

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2/rev x Q19761 ..  
Align seg 1/1 to: Q19761 from: 1 to: 2287

1636 AAGAAGATGCTGAATAATCAGAGGACT 1610  
|||||  
205 LysLysAspAlaGluAsnGlnArgThr 213

seq\_name: sp\_vertibrate:Q90YG4

seq\_documentation\_block:  
ID Q90YG4 PRELIMINARY; PRT; 77 AA.

AC Q90YG4;  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE PUTATIVE PHOSPHATIDYLETHANOLAMINE N-METHYLTRANSFERASE  
DE (FRAGMENT).  
GN PEMT.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pleguezuelos O., Secombes C.J.;  
RT "Screening a rainbow trout (Oncorhynchus mykiss) cDNA library."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ312343; CAC44161.1; -.  
KW Transferase; Methyltransferase.  
FT NON\_TER 77  
SQ SEQUENCE 77 AA; 8961 MW; 5BD75CF64E0281C5 CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x Q90YG4 ..  
Align seg 1/1 to: Q90YG4 from: 1 to: 77

1805 GCAGTAATAGCCATTATTTTCAT 1828  
|||||  
18 AlaValIleAlaIleIlePheAsn 25

seq\_name: sp\_virus:Q98158

seq\_documentation\_block:  
ID Q98158 PRELIMINARY; PRT; 95 AA.

AC Q98158; O12569;  
DT 01-FEB-1997 (TRENBLrel. 02, Created)

DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE ORF K6.  
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Rhadinovirus.  
OX NCBI\_TaxID=37296;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97094384; PubMed-8939871;  
RA Moore P.S., Bashoff C., Weiss R.A., Chang Y.;  
RT "Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV."  
RL Science 274:1739-1744(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97121480; PubMed-8962146;  
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,  
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;  
RT "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8)."  
RL proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,  
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Nicholas J., Ruvoilo V.R., Burns W.H., Sandford G., Wan X., Clufo D.,  
RA Hendrickson S., Guo H.G., Hayward G.S., Reitz M.S.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,  
RA Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97296220; PubMed-9151804;  
RA Neipel F., Albrecht J.C., Fleckenstein B.;  
RT "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus human herpesvirus 8: determinants of its pathogenicity?";  
RL J. Virol. 71:4187-4192(1997).  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Sun R., Lin S.-F., Miller G.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RA Ren S., Lin S.-F., Miller G.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U75698; AAC57095.1; -.  
DR EMBL; U74585; AAB61704.1; -.  
DR EMBL; U93872; AAB62671.1; -.  
DR EMBL; U71366; AAC34943.1; -.  
DR EMBL; U50138; AAD11536.1; -.  
DR HSSP; Q98157; LVMP.  
DR InterPro; IPR001811; Chemokine\_IL8.  
DR Pfam; PF00048; IL8; 1.  
DR SMART; SM00199; SCY; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 95 AA; 10485 MW; 34B9AFC4987FC485 CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2/rev x Q98158 ..

Align seg 1/1 to: Q98158 from: 1 to: 95

2792 GCTGGGCTCTTACGTGCATACACA 2769

|||||  
25 AlaGlySerLeuValSerIyrThr 32

seq\_name: sp\_mammal:Q9N0D9

seq\_documentation\_block:  
ID Q9N0D9 PRELIMINARY; PRT; 106 AA.  
AC Q9N0D9;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE UNNAMED PROTEIN PRODUCT.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
RA Suzuki Y., Sugano S., Hashimoto K.;  
RT "Isolation of full-length cDNA clones from macaque brain cDNA  
RT libraries."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB045991; BAB01573.1; -;  
SQ SEQUENCE 106 AA; 10952 MW; 311E1F0684EAE4EB CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x Q9N0D9 ..

Align seg 1/1 to: Q9N0D9 from: 1 to: 106

2846 TCCTCTGGGCTCTCCAGTTGGCC 2823

|||||  
9 SerLeuGlySerSerLeuAla 16

seq\_name: sp\_mammal:Q95JK7

seq\_documentation\_block:  
ID Q95JK7 PRELIMINARY; PRT; 106 AA.  
AC Q95JK7;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHETICAL 11.2 KDA PROTEIN.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TESTIS;  
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,  
RA Terao K., Sugano S.;  
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA  
RT libraries."  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB070174; BAB63119.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 106 AA; 11151 MW; 701E1F1B322C2230 CRC64;

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x Q95JK7 ..

Align seg 1/1 to: Q95JK7 from: 1 to: 106

2846 TCCTCTGGGCTCTCCAGTTGGCC 2823

|||||  
9 SerLeuGlySerSerLeuAla 16

seq\_name: sp\_bacteriap:O05260

seq\_documentation\_block:  
ID O05260 PRELIMINARY; PRT; 113 AA.  
AC O05260;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE HYPOTHETICAL 12.2 KDA PROTEIN.  
GN YUFV.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Oudega B., Koningssteyn G., Dueterhoeft A., Hilbert H.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelie D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pulic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
RT subtilis."  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z93937; CAB07944.1; -;

DR EMBL; Z99120; CAB15151.1; -.  
DR InterPro; IPR002757; DUF67.  
DR Pfam; PF01898; DUF67; 1.  
DR ProDom; PD006097; DUF67; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 113 AA; 12180 MW; 8681D1FF6F99112 CRC64;

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x O05260 ..

Align seg 1/1 to: O05260 from: 1 to: 113

630 CTGTTATTATCGAAATCCTTGTTA 653  
|||||  
19 LeuLeuSerLysSerLeuLeu 26

seq\_name: sp\_bacteria:Q9RPX2

seq\_documentation\_block:

ID Q9RPX2 PRELIMINARY; PRT; 117 AA.

AC Q9RPX2;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE H0318 HOMOLOG (FRAGMENT).

OS Haemophilus ducreyi.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=730;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=35000;  
RX MEDLINE=99426796; PubMed=10496878;  
RA Lewis D.A., Klesney-Tait J., Lumbley S.R., Ward C.K., Latimer J.L.,  
RA Ison C.A., Hansen E.J.;  
RT Identification of the znuA-encoded periplasmic zinc transport protein  
of Haemophilus ducreyi.;  
RL Infect. Immun. 67:5060-5068(1999).  
DR EMBL; AF141971; AAF00114.1; -.  
FT NON\_TER 117  
SQ SEQUENCE 117 AA; 13723 MW; F298B908B50869BF CRC64;

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9RPX2 ..

Align seg 1/1 to: Q9RPX2 from: 1 to: 117

2608 CTGTTCTCACTTTCATATCTGTA 2631  
|||||  
16 LeuPheSerLeuSerIleSerVal 23

seq\_name: sp\_vertebrate:Q9I8G7

seq\_documentation\_block:

ID Q9I8G7 PRELIMINARY; PRT; 135 AA.

AC Q9I8G7;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE VOLTAGE-DEPENDENT CALCIUM ION CHANNEL BETA SUBUNIT (FRAGMENT).  
OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INNER EAR;  
RA Fashena D.S., Trujillo-Provencio C., Serrano E.E.;  
RT "Cloning of calcium channels from Xenopus inner ear tissue."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases..  
DR EMBL; AF279368; AAF87794.1; -.  
DR InterPro; IPR000584; Ca\_channel\_B.  
DR InterPro; IPR000619; Guanylate\_kin.  
DR Pfam; PF00774; Ca\_channel\_B; 1.  
DR SMART; SM00072; GuKc; 1.  
FT NON\_TER 135  
FT NON\_TER 135  
SQ SEQUENCE 135 AA; 15114 MW; 9947511D52A806EE CRC64;

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x Q9I8G7 ..

Align seg 1/1 to: Q9I8G7 from: 1 to: 135

2746 CTGGTGATATTAGCGCTGATACC 2723  
|||||  
59 LeuValIleLeuAspAlaAspThr 66

seq\_name: sp\_bacteriap:Q930L0

seq\_documentation\_block:

ID Q930L0 PRELIMINARY; PRT; 144 AA.

AC Q930L0;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL PROTEIN SMA0343.  
GN SMA0343.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Plasmid pSymA (megaplasmid 1).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RX MEDLINE=21396509; PubMed=11481432;

RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
RT "Nucleotide sequence and predicted functions of the entire  
Sinorhizobium meliloti pSymA megaplasmid."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
DR EMBL; AE007211; AAK64843.1; -.  
KW Hypothetical protein; Plasmid; Complete proteome.  
SQ SEQUENCE 144 AA; 15219 MW; 30E8D8F3070FEA0D CRC64;

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x Q930L0 ..

Align seg 1/1 to: Q930L0 from: 1 to: 144

517 CCTCGCGTATCTCTAGTAGTGGT 494

|||||

61 ProAlaValSerLeuValAlaGly 68

seq\_name: sp\_bacteriap:Q97GU2

seq\_documentation\_block:

ID Q97GU2 PRELIMINARY; PRT; 149 AA.

AC Q97GU2;

DT 01-OCT-2001 (TREMBlrel. 18, Created)

DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE UNCHARACTERIZED PROTEIN YIHZ FAMILY.

GN CAC2273.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RT "Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum.";

RL J. Bacteriol. 183:4823-4838(2001).

DR EMBL: AE007728; AAK80230.1; "

DR InterPro: IPR003732; DUF154.

DR Pfam: PF02580; DUF154; 1.

KW Complete proteome.

SQ SEQUENCE 149 AA; 16586 MW; 69B03FDDC1A8791E CRC64;

alignment\_scores:

Quality: 8.00

Ratio: 1.000

Percent Similarity: 100.000

Length: 8

Gaps: 0

Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x Q97GU2 ..

Align seg 1/1 to: Q97GU2 from: 1 to: 149

773 AGTCTCTTAAGTAGATGCCAAA 750

|||||

11 SerSerValysValaspGlyLys 18

seq\_name: sp\_mammal:Q9MZA9

seq\_documentation\_block:

ID Q9MZA9 PRELIMINARY; PRT; 154 AA.

AC Q9MZA9;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE VIMENTIN (FRAGMENT).

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI\_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTAL ARTERY;

RA Zheng J., Tsol S.C., Magness R.R.;

RT "Growth factor expression in ovine fetal placental artery endothelial

RT cells.";

RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.

DR EMBL: AF251147; AAF87226.1; "

DR InterPro: IPR001664; IF.

DR Pfam: PF00038; filament; 1.

FT NON\_TER 154

SQ SEQUENCE 154 AA; 17171 MW; 3C1632E79DE31A49 CRC64;

alignment\_scores:

Quality: 8.00

Ratio: 1.000

Percent Similarity: 100.000

Length: 8

Gaps: 0

Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9MZA9 ..

Align seg 1/1 to: Q9MZA9 from: 1 to: 154

542 TCAGCCTTGAGGCCAACACATCT 565

|||||

42 SerAlaLeuArgProThrThrSer 49

seq\_name: sp\_bacteriap:Q9IOQ8

seq\_documentation\_block:

ID Q9IOQ8 PRELIMINARY; PRT; 186 AA.

AC Q9IOQ8;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE PROBABLE ACETYLTTRANSFERASE.

GN PA2578.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

RA Brody R.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";

RL Nature 406:959-964(2000).

DR EMBL: AE004686; AAG05966.1; "

DR InterPro: IPR000182; Acetyltransf\_GCN5.

DR Pfam: PF00583; Acetyltransf; 1.

KW Transferase; Complete proteome.

SQ SEQUENCE 186 AA; 21032 MW; 4FC6FA03706EB02A CRC64;

alignment\_scores:

Quality: 8.00

Ratio: 1.000

Percent Similarity: 100.000

Length: 8

Gaps: 0

Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x Q9IOQ8 ..

Align seg 1/1 to: Q9IOQ8 from: 1 to: 186

1029 ACTCCAGGAAGTGGATCTGTCCCA 1052

|||||

3 ThrProGlyThrGlySerValPro 10

seq\_name: sp\_invertebrate:O60965

seq\_documentation\_block:

```
ID O60965 PRELIMINARY; PRT; 189 AA.
AC O60965;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE L549.4.
GN L549.4.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RX MEDLINE=99178987; PubMed=10077609;
RA Myler P.J., Audleman L., devos T., Hixson G., Kiser P., Lemley C.,
RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of
RT protein-coding genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AE001274; AAC24616.1; -.
DR InterPro; IPR003169; GVF.
SQ SEQUENCE 189 AA; 20779 MW; 6897F4BA0C7C595D CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x O60965 ..
Align seg 1/1 to: O60965 from: 1 to: 189

3071 AGAAGTCCAAGTCTCCTCCAGCTCC 3094
|||||
31 ArgSerProSerSerSerSer 38

seq_name: sp_bacteria:Q93JB5

seq_documentation_block:
ID Q93JB5 PRELIMINARY; PRT; 191 AA.
AC Q93JB5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN SCBAC17A6.14C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL596248; CAC44659.1; -.
```

```
SQ SEQUENCE 191 AA; 20137 MW; 8F1B71DFE176B924 CRC64;

alignment_scores:
  Quality: 8.00 Length: 8
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x Q93JB5 ..
Align seg 1/1 to: Q93JB5 from: 1 to: 191

1932 CATGGTCTCTGCTGGTGAGGAAGA 1909
|||||
31 HisGlyProAlaGlyGlyGlyArg 38
```



OM of: US-09-202-054-2 to: Issued\_Patents\_AA:\* out\_format : pfs

Date: Jul 16, 2002 9:37 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh  
-Q/cgcn2\_1/JUSTO\_model/US09202054/runat\_16072002\_074912\_7743/app\_query.fasta\_1.3393  
-DB=Issued\_Patents\_AA -QFMT=fastan -SUFFIX=oli6.raii  
-GAPOP=4.500 -GAPOP=4.500 -GAPEXT=0.050 -LOOPLC=0.000  
-LOOPEXT=0.000 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000  
-YGAPEXT=60.000 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=oligo -TRANS=human40.cdi -LIST=1000 -DOCALIGN=200  
-THR SCORE=quality -THR\_MIN=6 -ALIGN=50 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09202054 -CGCN1\_L0 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-202-054-2

Query length: 3283

Database: Issued\_Patents\_AA:\*

Database sequences: 231628

Database length: 24425594

Search time (sec): 64.930000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000

WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score\_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-442-063A-42 +	9.00	132.74	2.80	236	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-442-063A-45 +	9.00	131.52	2.74	282	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-442-063B-48 +	9.00	130.93	2.71	307	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-07-613-083B-1 +	9.00	130.65	2.70	320	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-442-063A-27 +	9.00	130.38	2.69	333	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-272-919-2 +	9.00	130.19	2.68	342	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-619-916-2 +	9.00	130.19	2.68	342	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:PCT-US95-08542-2 +	9.00	130.19	2.68	342	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-303-238-4 +	9.00	129.86	2.66	359	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-458-834-4 +	9.00	129.86	2.66	359	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-07-821-717B-6 +	9.00	126.22	2.50	610	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-119-262B-6 +	9.00	126.22	2.50	610	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-135-929A-11 +	9.00	126.22	2.50	610	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-234-265A-11 +	9.00	126.22	2.50	610	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-09-353-585-8 +	8.00	124.13	30.69	65	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-09-230-371A-21 -	8.00	121.59	29.38	94	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-09-230-637-26 -	8.00	121.52	29.34	95	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-09-353-585-6 +	8.00	114.39	25.94	268	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-09-171-461-23 +	8.00	114.02	25.77	283	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-09-188-930-303 -	8.00	108.66	23.50	617	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-945-983-2 +	8.00	106.83	22.76	806	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-666-271-2 +	8.00	106.36	22.58	863	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-475-891A-4 +	8.00	105.26	22.15	1012	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-475-891A-2 +	8.00	105.19	22.13	1023	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-567-375-2 +	8.00	105.19	22.13	1023	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-587-680A-2 +	8.00	105.18	22.12	1023	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-567-375-4 +	8.00	105.18	22.12	1025	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-853-310-4 +	8.00	102.49	21.12	1514	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-07-942-245-494 -	7.00	118.86	356.47	11	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-07-664-989B-11 -	7.00	116.73	343.58	15	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-07-664-989B-18 -	7.00	116.73	343.58	15	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-268-251-24 -	7.00	116.73	343.58	15	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:PCT-US93-01112-24 -	7.00	116.73	343.58	15	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-268-251-54 -	7.00	113.79	326.57	23	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:PCT-US93-01112-54 -	7.00	113.79	326.57	23	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-486-013-17 -	7.00	109.18	301.56	45	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-482-273-17 -	7.00	109.18	301.56	45	i

/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-342-268-17 -	7.00	109.18	301.56	45	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-015-968-17 -	7.00	109.18	301.56	45	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-397-386-17 -	7.00	109.18	301.56	45	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:PCT-US92-08558-9 -	7.00	108.05	295.76	53	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-152-721B-28 -	7.00	105.40	285.67	71	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-130-339-4 -	7.00	105.40	282.49	78	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-710-749-13 +	7.00	103.76	274.61	99	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-710-749-14 +	7.00	103.76	274.61	99	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-710-749-16 +	7.00	103.76	274.61	99	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-340-203A-12 +	7.00	103.55	273.64	102	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-452-427-12 +	7.00	103.55	273.64	102	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-710-749-18 +	7.00	103.55	273.64	102	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-085-407-8 +	7.00	103.55	273.64	102	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-486-013-22 -	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-482-273-22 -	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-342-268-22 -	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-553-518B-3 +	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/6A_COMB.pcp:US-09-015-968-22 -	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-268-202-3 +	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-397-386-22 -	7.00	103.49	273.32	103	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-483-749A-28 -	7.00	103.10	271.49	109	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-07-942-245-1 -	7.00	101.94	266.11	129	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-943-136-2 -	7.00	101.94	266.11	129	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-973-518-2 -	7.00	101.53	264.22	137	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-258-026A-18 -	7.00	101.53	264.22	137	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:PCT-US95-07541-18 +	7.00	101.43	263.76	139	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-199-637A-101 +	7.00	101.00	261.81	148	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-193-877-5 +	7.00	100.72	260.57	154	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-706-214-4 +	7.00	100.59	259.98	157	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-350-486-90 +	7.00	100.59	259.98	157	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-228-986-90 +	7.00	100.46	259.39	160	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-117-257-38 +	7.00	100.46	259.39	160	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-489-352-38 +	7.00	99.96	257.18	172	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-923-738-2 +	7.00	99.96	257.18	172	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-923-738-4 +	7.00	99.92	257.18	172	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-936-165A-385 +	7.00	99.92	257.18	172	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-193-977-10 -	7.00	99.21	253.84	192	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-475-316A-23 +	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-233-005-4 -	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-428-943-4 -	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/6A_COMB.pcp:US-09-016-649-4 -	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-117-257-17 +	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-945-476-17 +	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-489-352-17 +	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:PCT-US95-04858-4 -	7.00	99.14	253.53	194	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-117-257-34 +	7.00	99.10	253.37	195	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-117-257-46 +	7.00	99.10	253.37	195	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-489-352-46 +	7.00	99.10	253.37	195	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-228-986-93 +	7.00	99.00	252.91	198	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-773-910-1 +	7.00	97.91	248.20	232	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-199-892-1 +	7.00	97.91	248.20	232	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-09-193-877-6 +	7.00	97.13	244.86	260	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-889-425-2 -	7.00	97.07	244.64	262	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-08-452-027A-4 +	7.00	96.82	243.56	272	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-409-452-2 -	7.00	96.74	243.24	275	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-188-930-145 -	7.00	96.12	240.64	301	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-188-930-149 -	7.00	96.12	240.64	301	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-045-973-1 +	7.00	96.07	240.45	303	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-757-553-172 -	7.00	95.57	238.37	326	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-823-516-75 +	7.00	95.57	238.37	326	i
/cgcn2_6/ptodata/2/iaa/5B_COMB.pcp:US-08-823-516-135 -	7.00	95.57	238.37	326	i
/cgcn2_6/ptodata/2/iaa/6A_COMB.pcp:US-08-759-038-111 -	7.00	95.57	238.37	326	i
/cgcn2_6/ptodata/2/iaa/6A_COMB.pcp:US-08-758-314-111 -	7.00	95.57	238.37	326	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-211-682-25 +	7.00	95.55	238.29	327	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-542-767A-1 +	7.00	94.89	235.58	360	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-091-405-2 +	7.00	94.63	234.52	374	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-525-697-2 +	7.00	94.57	234.30	377	i
/cgcn2_6/ptodata/2/iaa/5A_COMB.pcp:US-08-482-385A-6 +	7.00	94.50	234.00	381	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-416-050A-2 +	7.00	94.30	233.21	392	i
/cgcn2_6/ptodata/2/iaa/6B_COMB.pcp:US-09-665-309-2 +	7.00	94.30	233.21	392	i

[illegible]



/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-250-789A-73 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-248-839C-15 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-764-640-91 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-973-225-91 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-244-298A-91 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-101-167-11 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-07-516-704-91 - 6.00 99.73 3.3e+03 14 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-907-190-7 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-859-291C-36 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-314-172A-5 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-455-625-17 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-602-999A-304 + 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-455-665-17 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-060-988A-17 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US93-06751-3 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US93-06751-28 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US93-06751-35 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US94-05142-17 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US94-06655-4 - 6.00 99.25 3.2e+03 15 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-480-910-41 + 6.00 98.81 3.2e+03 16 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-488-379-41 + 6.00 98.81 3.2e+03 16 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-789-333F-15 + 6.00 98.81 3.2e+03 16 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-086-335C-1 - 6.00 98.39 3.2e+03 17 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-468-279-2 + 6.00 98.39 3.2e+03 17 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-889-425-5 + 6.00 98.39 3.2e+03 17 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-086-335C-3 - 6.00 98.00 3.2e+03 18 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-09-017-205-23 + 6.00 98.00 3.2e+03 18 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-09-227-357-376 - 6.00 98.00 3.2e+03 18 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-086-335C-8 - 6.00 97.63 3.1e+03 19 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-706-741B-30 - 6.00 97.63 3.1e+03 19 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-888-497-4 - 6.00 97.63 3.1e+03 19 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-924-695A-30 - 6.00 97.63 3.1e+03 19 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-09-362-230-4 + 6.00 97.63 3.1e+03 19 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US94-07926-4 + 6.00 97.63 3.1e+03 19 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-080-809-3 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-086-335C-2 - 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-339-696-107 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-484-530-47 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-827-618A-47 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-483-952A-47 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-256-747C-11 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-855-925A-2 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-834-130A-11 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-973-782-1 + 6.00 97.27 3.1e+03 20 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US96-11374-1 + 6.00 96.94 3.1e+03 21 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-086-335C-4 - 6.00 96.94 3.1e+03 21 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-169-015-43 - 6.00 96.94 3.1e+03 21 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-905-223-20 + 6.00 96.94 3.1e+03 21 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-924-030-2 + 6.00 96.62 3.1e+03 22 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-372-887-5 + 6.00 96.62 3.1e+03 22 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-976-358-24 + 6.00 96.02 3.1e+03 22 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-592-500-26 + 6.00 96.02 3.1e+03 24 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-195-006-26 + 6.00 96.02 3.1e+03 24 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US94-07644A-26 + 6.00 96.02 3.1e+03 24 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-976-358-21 + 6.00 95.74 3.0e+03 25 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-976-358-12 + 6.00 95.74 3.0e+03 25 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-969-931-10 - 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-855-147A-10 - 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-07-942-445-489 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-637-759B-332 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-871-355A-332 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-658-857B-13 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pcp:US-08-763-226C-13 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-09-307-200-13 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-09-201-945-332 + 6.00 95.47 3.0e+03 26 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-340-428B-12 + 6.00 94.96 3.0e+03 28 !  
/cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pcp:PCT-US93-07306-12 + 6.00 94.96 3.0e+03 28 !  
/cgn2\_6/ptodata/2/iaa/5A\_COMB.pcp:US-08-248-839C-134 - 6.00 94.72 3.0e+03 29 !  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-08-248-839C-136 - 6.00 94.72 3.0e+03 29 !  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pcp:US-08-248-839C-80 + 6.00 94.49 3.0e+03 30 !

/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-251-885-49 + 6.00 90.98 2.8e+03 50 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-337-635-43 + 6.00 90.98 2.8e+03 50 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-418-893D-3 - 6.00 90.84 2.8e+03 51 1 S  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-630-916A-117 + 6.00 90.84 2.8e+03 51 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-081-539-27 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-466-647-27 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-411-795B-146 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-411-795B-146 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-469-319A-146 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-471-039-146 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-851-843A-184 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-974-549A-303 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-854-050-184 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-430-323-184 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-277-078-46 + 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-584-043A-91 - 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-713-950-9 - 6.00 90.71 2.8e+03 52 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-314-268-151 + 6.00 90.45 2.8e+03 54 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-632-598-47 - 6.00 90.32 2.8e+03 55 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-177-249-139 - 6.00 90.32 2.8e+03 55 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-231-240-47 - 6.00 90.32 2.8e+03 55 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-227-357-553 + 6.00 90.32 2.8e+03 55 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-557-309B-29 - 6.00 90.20 2.8e+03 56 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-850-910A-24 + 6.00 90.20 2.8e+03 56 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-834-306A-29 - 6.00 90.20 2.8e+03 56 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-993-674A-29 - 6.00 90.20 2.8e+03 56 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-506-553C-1 - 6.00 90.08 2.8e+03 57 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-086-630C-28 - 6.00 89.96 2.8e+03 58 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-086-630C-43 - 6.00 89.96 2.8e+03 58 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-086-630C-219 - 6.00 89.96 2.8e+03 58 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-043-303-4 - 6.00 89.96 2.8e+03 58 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-248-839C-155 - 6.00 89.73 2.7e+03 60 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-936-165A-346 - 6.00 89.50 2.7e+03 62 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-450-360-10 + 6.00 89.07 2.7e+03 66 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-989-510A-5 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-182-816-5 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-182-816-20 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-471-528-5 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-471-528-20 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-634-530-5 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-634-530-20 + 6.00 88.67 2.7e+03 70 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-235-659-2 - 6.00 88.47 2.7e+03 72 1 S  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-338-579A-100 - 6.00 88.47 2.7e+03 72 1 S  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-378-579A-105 - 6.00 88.47 2.7e+03 72 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-905-223-372 + 6.00 88.47 2.7e+03 72 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-081-539-23 + 6.00 88.19 2.7e+03 75 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-466-647-23 + 6.00 88.19 2.7e+03 75 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-411-795B-139 + 6.00 88.19 2.7e+03 75 1 S  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-665-220-66 + 6.00 88.19 2.7e+03 75 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-469-319A-139 + 6.00 88.19 2.7e+03 75 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-231-692-66 + 6.00 88.19 2.7e+03 75 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-238-373-4 + 6.00 88.01 2.7e+03 77 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-726-306A-32 - 6.00 88.01 2.7e+03 77 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-227-357-552 + 6.00 88.01 2.7e+03 77 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-382-155-5 - 6.00 87.84 2.7e+03 79 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-382-155-9 - 6.00 87.84 2.7e+03 79 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-074-044A-5 - 6.00 87.84 2.7e+03 79 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-807-200-4 + 6.00 87.66 2.6e+03 81 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-001-777-4 + 6.00 87.66 2.6e+03 81 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-043-303-6 + 6.00 87.66 2.6e+03 81 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-382-155-11 + 6.00 87.50 2.6e+03 83 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-074-044A-11 + 6.00 87.50 2.6e+03 83 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-936-165A-410 - 6.00 87.50 2.6e+03 83 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-353-476-77 + 6.00 87.41 2.6e+03 84 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-679-493A-96 + 6.00 87.41 2.6e+03 84 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-480-229C-5 - 6.00 87.41 2.6e+03 84 1 Patent  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-480-229C-5 - 6.00 87.33 2.6e+03 85 1 S  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-659-235C-5 - 6.00 87.33 2.6e+03 85 1 S  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-103-478-27 + 6.00 87.33 2.6e+03 85 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-09-193-931C-27 + 6.00 87.25 2.6e+03 86 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-245-511-38 + 6.00 87.25 2.6e+03 86 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-933-750C-40 - 6.00 87.25 2.6e+03 86 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-600-993A-38 + 6.00 87.25 2.6e+03 86 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-09-234-613-40 - 6.00 87.25 2.6e+03 86 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-466-033-146 - 6.00 87.09 2.6e+03 88 1  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-444-733-146 - 6.00 87.09 2.6e+03 88 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-464-134-146 - 6.00 87.09 2.6e+03 88 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-461-361-146 - 6.00 87.09 2.6e+03 88 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-485-910-146 - 6.00 87.09 2.6e+03 88 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-816-977-8 - 6.00 87.02 2.6e+03 89 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-134-198E-4 - 6.00 86.94 2.6e+03 90 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-718-538-7 - 6.00 86.94 2.6e+03 90 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-09-231-529-9 - 6.00 86.79 2.6e+03 92 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-977-816-9 - 6.00 86.79 2.6e+03 92 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-486-580A-3 - 6.00 86.71 2.6e+03 93 1 Pat  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-486-013-21 + 6.00 86.50 2.6e+03 96 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-482-279-21 + 6.00 86.50 2.6e+03 96 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-342-268-21 + 6.00 86.50 2.6e+03 96 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-09-015-568-21 + 6.00 86.50 2.6e+03 96 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-397-386-21 + 6.00 86.50 2.6e+03 96 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-816-977-27 - 6.00 86.42 2.6e+03 99 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-09-047-125-23 + 6.00 86.28 2.6e+03 99 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-07-736-335E-23 + 6.00 86.28 2.6e+03 99 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-07-736-335E-29 - 6.00 86.28 2.6e+03 99 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-133-521-5 - 6.00 86.15 2.6e+03 101 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-09-374-135-6 + 6.00 86.08 2.6e+03 102 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-09-122-443-5 - 6.00 86.08 2.6e+03 102 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-09-023-221A-2 - 6.00 86.08 2.6e+03 102 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-282-352A-2 - 6.00 86.01 2.6e+03 103 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-905-223-473 + 6.00 85.95 2.6e+03 104 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-853-974-3 + 6.00 85.95 2.6e+03 104 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-172-988-3 + 6.00 85.88 2.6e+03 105 1  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-081-539-73 + 6.00 85.88 2.6e+03 105 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-466-647-73 + 6.00 85.88 2.6e+03 105 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-124-671-17 + 6.00 85.88 2.6e+03 105 1  
/cgn2\_6/ptodata/2/1aa/5B\_COMB.pcp:US-08-247-155-165 - 6.00 85.82 2.6e+03 106 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-557-309B-44 - 6.00 85.82 2.6e+03 106 1  
/cgn2\_6/ptodata/2/1aa/6A\_COMB.pcp:US-08-834-306-44 - 6.00 85.82 2.6e+03 106 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-083-351-15 - 6.00 85.82 2.6e+03 106 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-188-931-136 + 6.00 85.82 2.6e+03 106 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-083-352-15 - 6.00 85.82 2.6e+03 106 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-993-674A-44 - 6.00 85.82 2.6e+03 107 1  
/cgn2\_6/ptodata/2/1aa/5A\_COMB.pcp:US-08-359-796-17 + 6.00 85.75 2.6e+03 107 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-852-091-17 + 6.00 85.75 2.6e+03 107 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-478-097A-19 + 6.00 85.75 2.6e+03 107 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-08-939-853A-5 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-113-395-1 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-113-395-2 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-115-395-3 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-113-395-32 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-113-977C-32 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-113-977C-33 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-113-977C-34 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-507-102-1 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-507-102-2 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-507-102-3 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-250-059-1 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-250-059-2 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-250-059-3 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-248-074-2 - 6.00 85.69 2.6e+03 108 1  
/cgn2\_6/ptodata/2/1aa/6B\_COMB.pcp:US-09-248-074-2 - 6.00 85.69 2.6e+03 108 1



/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-13	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US96-09448-17	-	6.00	85.14	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-82	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US96-09448-18	-	6.00	85.14	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-83	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-300-386A-64	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-84	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-851-822A-3	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-259	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-948-612A-8	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-260	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-232-081B-8	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-261	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-09-193-510-8	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-262	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-09-184-384-3	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-263	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-965-904B-2	+	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-406	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-931-645-64	-	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-796-13	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-934-131-1	+	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-796-66	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-934-131-3	+	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-796-67	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-09-368-402-8	+	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-796-69	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-09-627-376-17	+	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-13	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-07213-13	+	6.00	85.08	2.5e+03	118
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-82	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US95-11235-64	-	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-83	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-497-025-6	+	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-84	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-561-521-12	-	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-259	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-561-521-13	-	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-260	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-665-259-34	+	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-261	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-762-500-34	+	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-262	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-09-012-515A-24	+	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-263	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-360-144A-24	+	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-406	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US95-01219-12	-	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-039-13	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US95-01219-13	-	6.00	85.02	2.5e+03	119
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-039-66	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-081-539-10	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-039-67	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-081-539-10	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-039-69	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-081-539-14	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-039-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-466-647-10	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-469-318-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-466-647-10	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-469-318-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-466-647-14	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-469-318-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-466-647-14	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-468-609A-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-468-609A-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-468-609A-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-402	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-192-325B-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-795B-402	-	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-192-325B-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-796-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-192-325B-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-411-796-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-045-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-045-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-045-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-402	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-09-023-221A-20	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-469-319A-402	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-469-712A-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-851-822A-1	-	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-469-712A-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-09-184-384-1	-	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-469-712A-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-471-039-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-193-373A-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-471-039-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-193-373A-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-09-056-536-195	-	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-193-373A-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-09-056-536-195	-	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-09-282-352A-20	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-08-871-488A-19	-	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-13	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-135	+	6.00	84.96	2.5e+03	120
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-66	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-478-039-64	-	6.00	84.91	2.5e+03	121
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-67	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-478-039-64	-	6.00	84.91	2.5e+03	121
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-69	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-478-039-64	-	6.00	84.91	2.5e+03	121
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-25	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-09-171-982-2	-	6.00	84.91	2.5e+03	121
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-26	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-561-521-11	-	6.00	84.79	2.5e+03	123
/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US93-11198-27	+	6.00	85.38	2.5e+03	113	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-09-448-280-1	-	6.00	84.79	2.5e+03	123
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-081-539B-64	+	6.00	85.32	2.5e+03	114	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US95-01219-11	-	6.00	84.79	2.5e+03	123
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-466-647-64	+	6.00	85.32	2.5e+03	114	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-455-559-11	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-479-025-8	+	6.00	85.32	2.5e+03	114	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-891-848-17	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-040-548-4	+	6.00	85.32	2.5e+03	114	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-875-811-12	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-466-344-4	+	6.00	85.32	2.5e+03	114	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-09-145-060-11	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-209-747-17	+	6.00	85.26	2.5e+03	115	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-09-192-242E-3	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-458-298-17	-	6.00	85.26	2.5e+03	115	1	/cgn2_6/ptodata/2/iaa/6B COMB. pep: US-09-199-242E-5	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-687-895-3	-	6.00	85.20	2.5e+03	116	1	/cgn2_6/ptodata/2/iaa/PCTUS_COMB. pep: PCT-US94-00657-11	+	6.00	84.74	2.5e+03	124
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-816-241-3	-	6.00	85.20	2.5e+03	116	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-275-370-23	+	6.00	84.68	2.5e+03	125
/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-09-040-482-3	-	6.00	85.20	2.5e+03	116	1	/cgn2_6/ptodata/2/iaa/5A COMB. pep: US-08-367-968-23	+	6.00	84.68	2.5e+03	125
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-09-128-395-3	-	6.00	85.20	2.5e+03	116	1	/cgn2_6/ptodata/2/iaa/5B COMB. pep: US-08-473-981A-10	+	6.00	84.68	2.5e+03	125
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-290-592E-17	-	6.00	85.14	2.5e+03	117	1	/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-665-484-23	+	6.00	84.68	2.5e+03	125
/cgn2_6/ptodata/2/iaa/6A COMB. pep: US-08-290-592E-18	-	6.00	85.14	2.5e+03	117	1	/cgn					







```
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: PIERSCHBACHER, MICHAEL D.
; APPLICANT: CARDENAS, JOSE
; APPLICANT: CRAIG, WILLIAM
; APPLICANT: MULLEN, DANIEL G.
; TITLE OF INVENTION: DECORIN FRAGMENTS AND METHODS OF
; TITLE OF INVENTION: INHIBITING CELL REGULATORY FACTORS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,063A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/865,652
; FILING DATE: 03-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1454
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 307 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-442-063A-48

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-08-442-063A-48 ..
Align seg 1/1 to: US-08-442-063A-48 from: 1 to: 307

508 ATACCGCAGGCGCTCCGCGCTAGCTTA 534
|||||
187 ileProGlnGlyLeuProSerLeu 195

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-613-083B-1
seq_documentation_block:
; Sequence 1, Application US/07613083B
; Patent No. 5340727
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Zaverio M.
; APPLICANT: Ware, Jerry, inventors
; APPLICANT: on behalf of Scripps Clinic and Research
; APPLICANT: Foundation
; TITLE OF INVENTION: GPIb' Fragments and Recombinant
; TITLE OF INVENTION: DNA Expression Vectors
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scripps Clinic and Research
; ADDRESSEE: Foundation
```

```
; STREET: 10666 No. 5340727th Torrey Pines Road
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb
; COMPUTER: AST Bravo IBM PC comp. (386SX)
; OPERATING SYSTEM: MS DOS version 3.2
; SOFTWARE: WordPerfect 5.1 conv. to ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/613,083B
; FILING DATE: 19911114
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: This appl. is a c-1-p of
; APPLICATION NUMBER: U.S. 07/470,674
; FILING DATE: 04-Jan-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barron, Alexis
; REGISTRATION NUMBER: 22,702
; REFERENCE/DOCKET NUMBER: P16,569-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 923-4466
; TELEFAX: (215) 923-2189
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320
; TYPE: AMINO ACID
; STRANDEDNESS: No. 5340727 applicable
; TOPOLOGY: Linear
; US-07-613-083B-1

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-07-613-083B-1 ..
Align seg 1/1 to: US-07-613-083B-1 from: 1 to: 320

2191 ACTTTGGACCTCAGCCACCAACCAACTG 2217
|||||
81 ThrLeuAspLeuSerHisAsnGlnLeu 89

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-442-063A-27
seq_documentation_block:
; Sequence 27, Application US/08442063A
; Patent No. 5705609
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: PIERSCHBACHER, MICHAEL D.
; APPLICANT: CARDENAS, JOSE
; APPLICANT: CRAIG, WILLIAM
; APPLICANT: MULLEN, DANIEL G.
; TITLE OF INVENTION: DECORIN FRAGMENTS AND METHODS OF
; TITLE OF INVENTION: INHIBITING CELL REGULATORY FACTORS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/442,063A  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/865,652  
; FILING DATE: 03-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 1454  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-535-9001  
; TELEFAX: 619-535-8949  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 333 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-442-063A-27

alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x US-08-442-063A-27 ..

Align seg 1/1 to: US-08-442-063A-27 from: 1 to: 333

508 ATACCGCAGGCGCTCCGGCTAGCTTA 534  
|||||  
187 IleProGlnGlyLeuProSerLeu 195

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-272-919-2

seq\_documentation\_block:  
; Sequence 2, Application US/08272919  
; Patent No. 5567807  
; GENERAL INFORMATION:  
; APPLICANT: Craig, William S.  
; APPLICANT: Harper, John R.  
; APPLICANT: Hernandez, Sam D.  
; APPLICANT: Kostel, Paul J.  
; APPLICANT: Parker, Jonathan R.  
; APPLICANT: Vedvick, Thomas S.  
; TITLE OF INVENTION: Processes for the Purification of Human  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/272,919  
; FILING DATE: 08-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 1040  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 342 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-272-919-2  
  
alignment\_scores:  
Quality: 9.00 Length: 9  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x US-08-272-919-2 ..  
  
Align seg 1/1 to: US-08-272-919-2 from: 1 to: 342  
  
508 ATACCGCAGGCGCTCCGGCTAGCTTA 534  
|||||  
199 IleProGlnGlyLeuProSerLeu 207  
  
seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-619-916-2  
  
seq\_documentation\_block:  
; Sequence 2, Application US/08619916  
; Patent No. 5763276  
; GENERAL INFORMATION:  
; APPLICANT: Craig, William S.  
; APPLICANT: Harper, John R.  
; APPLICANT: Hernandez, Sam D.  
; APPLICANT: Kostel, Paul J.  
; APPLICANT: Parker, Jonathan R.  
; APPLICANT: Vedvick, Thomas S.  
; TITLE OF INVENTION: Processes for the Purification of Human  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/619,916  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/272,919  
; FILING DATE: 08-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 1040  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 342 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-619-916-2



```
seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-303-238-4
seq_documentation_block:
; Sequence 4, Application US/08303238
; Patent No. 5654270
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: LONGAKER, MICHAEL T.
; APPLICANT: WHITBY, DAVID J.
; APPLICANT: HARPER, JOHN R.
; APPLICANT: PIERSCHBACHER, MICHAEL D.
; APPLICANT: BORDER, WAYNE A.
; TITLE OF INVENTION: INHIBITORS OF CELL REGULATORY FACTORS
; TITLE OF INVENTION: AND METHODS FOR PREVENTING OR REDUCING SCARRING
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,238
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/978,931
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KONSKI, ANTOINETTE F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: P-LA 9453
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-303-238-4

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-08-619-916-2 ..
Align seg 1/1 to: US-08-619-916-2 from: 1 to: 342
508 ATACCGAGGCGCTCCGCTAGCTTA 534
|||||
199 IleProGlnGlyLeuProSerLeu 207

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:PCT-US95-08542-2
seq_documentation_block:
; Sequence 2, Application PC/TUS9508542
; GENERAL INFORMATION:
; APPLICANT: La Jolla Cancer Research Foundation
; TITLE OF INVENTION: Processes for the Purification of
; TITLE OF INVENTION: Human
; TITLE OF INVENTION: Recombinant Decorin and the Detection of Guanidinium
; TITLE OF INVENTION: Ions
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08542
; FILING DATE: 07-JUL-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Susan M.
; REGISTRATION NUMBER: 36,405
; REFERENCE/DOCKET NUMBER: FP-LA 1751
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08542-2

alignment_scores:
Quality: 9.00 Length: 9
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x PCT-US95-08542-2 ..
Align seg 1/1 to: PCT-US95-08542-2 from: 1 to: 342
508 ATACCGAGGCGCTCCGCTAGCTTA 534
|||||
199 IleProGlnGlyLeuProSerLeu 207

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-458-834-4
seq_documentation_block:
; Sequence 4, Application US/08458834
; Patent No. 6277812
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: LONGAKER, MICHAEL T.
; APPLICANT: WHITBY, DAVID J.
```

```
;
; APPLICANT: HARPER, JOHN R.
; APPLICANT: PIERSCHBACHER, MICHAEL D.
; APPLICANT: BORDER, WAYNE A.
; TITLE OF INVENTION: INHIBITORS OF CELL REGULATORY FACTORS
; TITLE OF INVENTION: AND METHODS FOR PREVENTING OR REDUCING SCARRING
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,834
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/978,931
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOSKI, ANTOINETTE F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: P-LA 9453
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-458-834-4

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-08-458-834-4 ..
Align seg 1/1 to: US-08-458-834-4 from: 1 to: 359
508 ATACCGAGGCGCTCCGCTAGCTTA 534
215 IleProGlnGlyLeuProSerLeu 223

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:us-07-821-717B-6

seq_documentation_block:
; Sequence 6, Application US/07821717B
; Patent No. 5298239
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Cunningham, David
; APPLICANT: Lyle, Vicki A.
; APPLICANT: Finch, Clara N.
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Nixon, Harrgrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
```

```
;
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,717B
; FILING DATE: 15-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Timain, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20884/21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PUBLICATION INFORMATION:
; AUTHORS: Lopez, Jose A.
; AUTHORS: Chung, Dominic W.
; AUTHORS: Fujikawa, Kazuo
; AUTHORS: Hagen, Frederick S.
; AUTHORS: Papayannopoulou, Thalia
; AUTHORS: Roth, Gerald J.
; TITLE: Cloning of the alpha chain of human
; TITLE: platelet glycoprotein Ib: A transmembrane protein with homology
; TITLE: to leucine-rich alpha-2-glycoprotein
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 84
; PAGES: 5615-5619
; DATE: AUG-1987
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610
; PUBLICATION INFORMATION:
; AUTHORS: Zimmerman, Theodore S.
; AUTHORS: Ruggeri, Zaverio M.
; AUTHORS: Houghten, Richard A.
; AUTHORS: Vincete, Vincete
; AUTHORS: Mohri, Hiroshi
; TITLE: Proteolytic fragments and synthetic
; TITLE: peptides that block the binding of von Willebrand factor to the
; TITLE: platelet membrane glycoprotein Ib
; DOCUMENT NUMBER: EP 0 317 278 A2
; FILING DATE: 16-NOV-1988
; PUBLICATION DATE: 24-MAY-1989
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293
; US-07-821-717B-6

alignment_scores:
  Quality: 9.00 Length: 9
  Ratio: 1.000 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-07-821-717B-6 ..
Align seg 1/1 to: US-07-821-717B-6 from: 1 to: 610
2191 ACTTTGGACCTCAGCCACCAACCACTG 2217
81 ThrLeuAspLeuSerHisnGlnLeu 89

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-119-262B-6

seq_documentation_block:
```

Sequence 6, Application US/08119262B  
Patent No. 5492809  
GENERAL INFORMATION:  
APPLICANT: Miller, Jonathan L.  
APPLICANT: Cunningham, David  
APPLICANT: Lyle, Vicki A.  
APPLICANT: Finch, Clara N.  
TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,262B  
FILING DATE: 09-SEP-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER: US 07/821,717  
FILING DATE: 15-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Timain, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PUBLICATION INFORMATION:  
AUTHORS: Lopez, Jose A.  
AUTHORS: Chung, Dominic W.  
AUTHORS: Fujikawa, Kazuo  
AUTHORS: Hagen, Frederick S.  
AUTHORS: Papayannopoulou, Thalia  
AUTHORS: Roth, Gerald J.  
TITLE: Cloning of the alpha chain of human platelet  
TITLE: glycoprotein Ib: A transmembrane protein  
TITLE: leucine-rich alpha-2-glycoprotein  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 84  
PAGES: 5615-5619  
DATE: AUG-1987  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
PUBLICATION INFORMATION:  
AUTHORS: Zimmerman, Theodore S.  
AUTHORS: Ruggeri, Zaverio M.  
AUTHORS: Houghten, Richard A.  
AUTHORS: Vincete, Vincete  
AUTHORS: Mohri, Hiroshi  
TITLE: Proteolytic fragments and synthetic peptides  
TITLE: that block the binding of von Willebrand  
TITLE: membrane glycoprotein Ib  
DOCUMENT NUMBER: EP 0 317 278 A2  
FILING DATE: 16-NOV-1988  
PUBLICATION DATE: 24-MAY-1989  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-08-119-262B-6

alignment\_scores:                   Quality:   9.00           Length:    9  
                                  Ratio:    1.000           Gaps:     0  
Percent Similarity: 100.000   Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x US-08-119-262B-6   ..  
Align seg 1/1   to: US-08-119-262B-6   from: 1   to: 610  
2191 ACTTTGGACCTCAGCCACCAACCACTG 2217  
|||||  
81 ThrLeuAspLeuSerHisAsnGlnLeu 89  
seq\_name: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:US-08-135-929A-11  
seq\_documentation\_block:  
; Sequence 11, Application US/08135929A  
; Patent No. 5593959  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; APPLICANT: Pincus, Matthew R.  
; TITLE OF INVENTION: Mutations in the Gene Encoding the Alpha  
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/135,929A  
; FILING DATE: 14-OCT-1993  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/23  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; TELEX: 978450  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-135-929A-11  
alignment\_scores:                   Quality:   9.00           Length:    9  
                                  Ratio:    1.000           Gaps:     0  
Percent Similarity: 100.000   Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x US-08-135-929A-11   ..  
Align seg 1/1   to: US-08-135-929A-11   from: 1   to: 610  
2191 ACTTTGGACCTCAGCCACCAACCACTG 2217

with homology

factor to the

```
|||||
81 ThrLeuAspLeuSerHisAsnGlnLeu 89

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-234-265A-11

seq_documentation_block:
; Sequence 11, Application US/08234265A
; Patent No. 5624817
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Cunningham, David
; APPLICANT: Lyle, Vicki A.
; APPLICANT: Finch, Clara N.
; APPLICANT: Pincus, Matthew R.
; TITLE OF INVENTION: Mutations in the Gene Encoding the
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/234,265A
; FILING DATE: 28-APR-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Timlan, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20884/24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; TELEX: 978450
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-234-265A-11
```

```
alignment_scores:
  Quality: 9.00      Length: 9
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-202-054-2 x US-08-234-265A-11 ..
```

```
Align seg 1/1 to: US-08-234-265A-11 from: 1 to: 610
```

```
2191 ACTTGGACCTCAGCCACACCAACTG 2217
```

```
|||||
81 ThrLeuAspLeuSerHisAsnGlnLeu 89
```

```
seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-353-585-8
```

```
seq_documentation_block:
; Sequence 8, Application US/09353585
; Patent No. 6287865
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
```

```
;
;
; TITLE OF INVENTION: Plant pathogen resistance genes and uses
; thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6287865th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/353,585
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12O
; 1/68
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/930,277
; FILING DATE: 27-OCT-1997
; APPLICATION NUMBER: PCT/GB96/00785
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: GB 9506658.5
; FILING DATE: 31-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-69
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 65 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: Cf9
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-353-585-8
```

```
alignment_scores:
  Quality: 8.00      Length: 8
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-202-054-2 x US-09-353-585-8 ..
```

```
Align seg 1/1 to: US-09-353-585-8 from: 1 to: 65
```

```
2338 CTGGATCTCAGCTCAATAAATC 2361
```

```
|||||
1 LeuAspLeuSerSerAsnLysIle 8
```

```
seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-230-371A-21
```

```
seq_documentation_block:
; Sequence 21, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
```

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.ppt:US-09-353-585-6

seq\_documentation\_block:

Sequence 6, Application US/09353585

Patent No. 6287865

GENERAL INFORMATION:

APPLICANT: Dixon, Mark S

Jones, Jonathan DG

TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderhye PC

STREET: 8th Floor, 1100 No. 6287865th Glebe Road

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/353,585

FILING DATE: 15-Jul-1999

CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q 1/68

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/930,277

FILING DATE: 27-OCT-1997

APPLICATION NUMBER: PCT/GB96/00785

FILING DATE: 01-APR-1996

APPLICATION NUMBER: GB 9506658.5

FILING DATE: 31-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Ms Mary J Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 620-69

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: YES

FRAGMENT TYPE: C-terminal

ORIGINAL SOURCE:

ORGANISM: Tomato

STRAIN: Cf9

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-353-585-6

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-09-230-371A-21 ..

Align seg 1/1 to: US-09-230-371A-21 from: 1 to: 94

2792 GCTGGGCTCTTTAGTGTACATACACA 2769

25 AlaGlySerLeuValSerTyrThr 32

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.ppt:US-09-230-637-26

seq\_documentation\_block:

Sequence 26, Application US/09230637

Patent No. 6264958

GENERAL INFORMATION:

APPLICANT: Hayward, Gary

APPLICANT: Nicholas, John

APPLICANT: Hardwick, J. Marie

APPLICANT: Reitz, Marvin

TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma

TITLE OF INVENTION: Associated Herpesvirus

FILE REFERENCE: 1107.78372

CURRENT APPLICATION NUMBER: US/09/230,637

CURRENT FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: 60/022,591

PRIOR FILING DATE: 1996-07-25

PRIOR APPLICATION NUMBER: PCT US 97/12931

PRIOR FILING DATE: 1997-07-24

NUMBER OF SEQ ID NOS: 62

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 26

LENGTH: 95

TYPE: PRT

ORGANISM: Kaposi's sarcoma-associated herpes-like virus

US-09-230-637-26

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-09-230-637-26 ..

Align seg 1/1 to: US-09-230-637-26 from: 1 to: 95

2792 GCTGGGCTCTTTAGTGTACATACACA 2769

25 AlaGlySerLeuValSerTyrThr 32

seq\_documentation\_block:

Sequence 6, Application US/09353585

Patent No. 6287865

GENERAL INFORMATION:

APPLICANT: Dixon, Mark S

Jones, Jonathan DG

TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderhye PC

STREET: 8th Floor, 1100 No. 6287865th Glebe Road

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/353,585

FILING DATE: 15-Jul-1999

CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q 1/68

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/930,277

FILING DATE: 27-OCT-1997

APPLICATION NUMBER: PCT/GB96/00785

FILING DATE: 01-APR-1996

APPLICATION NUMBER: GB 9506658.5

FILING DATE: 31-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Ms Mary J Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 620-69

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: YES

FRAGMENT TYPE: C-terminal

ORIGINAL SOURCE:

ORGANISM: Tomato

STRAIN: Cf9

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-353-585-6

alignment\_scores:

Quality: 8.00 Length: 8

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-09-353-585-6 ..

Align seg 1/1 to: US-09-353-585-6 from: 1 to: 268

2338 CTGGATCTCAGCTCAATAAATC 2361

126 LeuAspLeuSerAsnLysIle 133

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-09-171-461-23

seq\_documentation\_block:

; Sequence 23, Application US/09171461  
; Patent No. 6335016  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Adam  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Chiocca, Susanna  
; APPLICANT: Kurbauer, Robert  
; APPLICANT: Schaffner, Gothold  
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus  
; FILE REFERENCE: 0652.1800000  
; CURRENT APPLICATION NUMBER: US/09/171.461  
; EARLIER FILING DATE: 1999-01-12  
; EARLIER APPLICATION NUMBER: PCT/EP97/01944  
; EARLIER FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: CELO Virus  
; FEATURE:  
; OTHER INFORMATION: Position: 41002..41853 /note=OREF10  
US-09-171-461-23

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-09-171-461-23 ..

Align seg 1/1 to: US-09-171-461-23 from: 1 to: 283

286 ACGAACCTCACCTCACCATTAAAC 309

|||||

222 ThrAsnLeuThrLeuThrIleAsn 229

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-09-188-930-303

seq\_documentation\_block:

; Sequence 303, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murlison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188.930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 303  
; LENGTH: 617  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-188-930-303

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-09-188-930-303 ..

Align seg 1/1 to: US-09-188-930-303 from: 1 to: 617

1920 GTGGAGGAAGAGATGTCATTGTC 1897

|||||

457 GlyGlyGlyArgAspValIleVal 464

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-08-945-983-2

seq\_documentation\_block:

; Sequence 2, Application US/08945983  
; Patent No. 6225527  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Colwyn M  
; APPLICANT: Ballint-Kurti, Peter J  
; APPLICANT: Jones, David A  
; APPLICANT: Jones, Jonathan DG  
; TITLE OF INVENTION: Plant pathogen resistance genes and uses  
; TITLE OF INVENTION: thereof  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye PC  
; STREET: 8th Floor, 1100 No. 6225527th Glebe Road  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/945.983  
; FILING DATE: 12-NOV-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB96/01155  
; FILING DATE: 13-MAY-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9509575.8  
; FILING DATE: 11-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ms Mary J Wilson  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-27  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 806 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: lycopersicon hirsutum  
; STRAIN: Cf4  
US-08-945-983-2

alignment\_scores:

Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-08-945-983-2 ..

Align seg 1/1 to: US-08-945-983-2 from: 1 to: 806

```
2338 CTGGATCTCAGCTCAATAAATC 2361
|||||
664 LeuAspLeuSerAsnLysile 671

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.US-08-666-271-2

seq_documentation_block:
; Sequence 2, Application US/08666271
; Patent No. 5920000
; GENERAL INFORMATION:
; APPLICANT: JONES, JONATHAN D
; APPLICANT: HAMMOND-KOSACK, KIM E
; APPLICANT: THOMAS, COLWYN M
; APPLICANT: JONES, DAVID A
; TITLE OF INVENTION: PLANT PATHOGEN RESISTANCE GENES AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHIE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,271
; FILING DATE: 19-SEP-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/02812
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326428.1
; FILING DATE: 24-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9409363.0
; FILING DATE: 11-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 620-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4091
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 863 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-666-271-2

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-08-666-271-2 ..
Align seg 1/1 to: US-08-666-271-2 from: 1 to: 863

2338 CTGGATCTCAGCTCAATAAATC 2361
|||||
721 LeuAspLeuSerAsnLysile 728

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.US-08-475-891A-4

seq_documentation_block:
; Sequence 4, Application US/08475891A
; Patent No. 5859339
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,891A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 02370-058910US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1012
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp.
; OTHER INFORMATION: disease resistance polypeptide RRR-B
; OTHER INFORMATION: from rice (Oryza sativa)"
; US-08-475-891A-4

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2 x US-08-475-891A-4 ..
Align seg 1/1 to: US-08-475-891A-4 from: 1 to: 1012

2194 TTGGACCTCAGCCACACCACTG 2217
|||||
156 LeuAspLeuSerHisAsnGlnLeu 163

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.US-08-475-891A-2

seq_documentation_block:
; Sequence 2, Application US/08475891A
; Patent No. 5859339
```

; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; TITLE OF INVENTION: Procedures and Materials for Conferring  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,891A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 02370-058910US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1023 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..1023  
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp.  
; OTHER INFORMATION: disease resistance polypeptide RRK-F  
; OTHER INFORMATION: from rice (Oryza sativa)"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1010  
; OTHER INFORMATION: /note= "Xaa = Leu when position  
; OTHER INFORMATION: 5471 of RRK-F = G or Phe when position  
; OTHER INFORMATION: 5471 of RRK-F = C"  
; US-08-475-891A-2

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x US-08-475-891A-2 ..

Align seg 1/1 to: US-08-475-891A-2 from: 1 to: 1023

2194 TTGACCTCAGCCACACCAACTG 2217  
|||||  
166 LeuAspLeuSerHisAsnGlnLeu 173

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-567-375-2

seq\_documentation\_block:

; Sequence 2, Application US/08567375

; Patent NO. 5952485  
; GENERAL INFORMATION:  
; APPLICANT: Ronald, Pamela C.  
; APPLICANT: Wang, Guo-Liang  
; APPLICANT: Song, Wen-Yuang  
; APPLICANT: Szabo, Veronique  
; TITLE OF INVENTION: Procedures and Materials for Conferring  
; TITLE OF INVENTION: Disease Resistance in Plants  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/567,375  
; FILING DATE: 04-DEC-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/004,645  
; FILING DATE: 29-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/475,891  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,375  
; FILING DATE: 17-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 023070-058930  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1023 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..1023  
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease  
; OTHER INFORMATION: resistance polypeptide RRK-F from rice  
; OTHER INFORMATION: (Oryza sativa)"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1010  
; OTHER INFORMATION: /product= "OTHER"  
; OTHER INFORMATION: /note= "Xaa = Leu when position 5471 of  
; OTHER INFORMATION: RRK-F = G or Phe when position 5471 of  
; OTHER INFORMATION: RRK-F = C"  
; US-08-567-375-2

alignment\_scores:  
Quality: 8.00 Length: 8  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x US-08-567-375-2 ..

Align seg 1/1 to: US-08-567-375-2 from: 1 to: 1023



```
2194 TTGGACCTCAGCCACACCACTG 2217
|||||
166 LeuAspLeuSerHisAsnGlnLeu 173

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.US-08-587-680A-2

seq_documentation_block:
; Sequence 2, Application US/08587680A
; Patent No. 5977434
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,680A
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/567,375
; FILING DATE: 04-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058940US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1023
; OTHER INFORMATION: /note= "Xa21 Xanthomonas spp. disease
; OTHER INFORMATION: resistance polypeptide RRK-F from rice
; OTHER INFORMATION: (Oryza sativa)"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1010
; OTHER INFORMATION: /note= "Xaa - Ieu when position 5471 of
; OTHER INFORMATION: RRK-F = G or Phe when position 5471 of
; OTHER INFORMATION: RRK-F = C"
US-08-587-680A-2
```

```
alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-202-054-2 x US-08-587-680A-2 ..
Align seg 1/1 to: US-08-587-680A-2 from: 1 to: 1023

2194 TTGGACCTCAGCCACACCACTG 2217
|||||
166 LeuAspLeuSerHisAsnGlnLeu 173

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.US-08-567-375-4

seq_documentation_block:
; Sequence 4, Application US/08567375
; Patent No. 5952485
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Szabo, Veronique
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,375
; FILING DATE: 04-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,645
; FILING DATE: 29-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,891
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,375
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058930
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: protein
US-08-567-375-4

alignment_scores:
Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-08-567-375-4 ..

Align seg 1/1 to: US-08-567-375-4 from: 1 to: 1025

2194 TTGGACCTCAGCCACCACTG 2217

|||||

156 LeuAspLeuSerHisAsnGlnLeu 163

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-587-680A-4

seq\_documentation\_block:

; Sequence 4, Application US/08587680A

; Patent No. 5977434

; GENERAL INFORMATION:

; APPLICANT: Ronald, Pamela C.

; APPLICANT: Wang, Guo-Liang

; APPLICANT: Song, Wen-Yuang

; APPLICANT: Szabo, Veronique

; TITLE OF INVENTION: Procedures and Materials for Conferring

; TITLE OF INVENTION: Disease Resistance in Plants

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/587.680A

; FILING DATE: 17-JAN-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/373,375

; FILING DATE: 17-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/475,891

; FILING DATE: 07-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/004,645

; FILING DATE: 29-SEP-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/567,375

; FILING DATE: 04-DEC-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Bastian, Kevin L.

; REGISTRATION NUMBER: 34,774

; REFERENCE/DOCKET NUMBER: 023070-058940US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1025 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-587-680A-4

alignment\_scores:

Quality: 8.00

Ratio: 1.000

Percent Similarity: 100.000 Percent Identity: 100.000

Length: 8

Gaps: 0

alignment\_block:

US-09-202-054-2 x US-08-587-680A-4 ..

Align seg 1/1 to: US-08-587-680A-4 from: 1 to: 1025

2194 TTGGACCTCAGCCACCACTG 2217

|||||

156 LeuAspLeuSerHisAsnGlnLeu 163

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-853-310-4

seq\_documentation\_block:

; Sequence 4, Application US/08853310

; Patent No. 5948640

; GENERAL INFORMATION:

; APPLICANT: Randazzo, Filippo

; TITLE OF INVENTION: Mammalian Additional Sex Combs (Asx) Acts as a Tumor Suppre

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 94608

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/853.310

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Guth, Joseph H.

; REGISTRATION NUMBER: 31,261

; REFERENCE/DOCKET NUMBER: 1228.003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (510) 923-3888

; TELEFAX: (510) 655-3542

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1514 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-853-310-4

alignment\_scores:

Quality: 8.00

Ratio: 1.000

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-08-853-310-4 ..

Align seg 1/1 to: US-08-853-310-4 from: 1 to: 1514

3074 AGTCCAAGTTCTCCAGCTCCGGA 3097

|||||

201 SerProSerSerSerSerGly 208

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-942-245-494

seq\_documentation\_block:

; Sequence 494, Application US/07942245

; Patent No. 5639641

; GENERAL INFORMATION:

; APPLICANT: PEDERSEN, Jan T.

; APPLICANT: SEARLE, Stephen M.J.

; APPLICANT: REES, Anthony R.  
; APPLICANT: ROGUSKA, Michael A.  
; APPLICANT: GUILD, Braydon C.  
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT  
; TITLE OF INVENTION: ANTIBODIES  
; NUMBER OF SEQUENCES: 522  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: United States  
; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: HP 9000/700 Workstation  
; OPERATING SYSTEM: UNIX  
; SOFTWARE: In house  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,245  
; FILING DATE: 09-SEP-1992  
; CLASSIFICATION: 530  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 293-7060  
; TELEFAX: (202) 293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 494:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-942-245-494

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-07-942-245-494 ..

Align seg 1/1 to: US-07-942-245-494 from: 1 to: 11

3067 AGGGCTTCTCAAGAAATATCA 3047

|||||  
1 ArgAlaSerGlnGluLeuSer 7

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-664-989B-11

seq\_documentation\_block:  
; Sequence 11, Application US/07664989B  
; Patent No. 5223409  
; GENERAL INFORMATION:  
; APPLICANT: Ladner, Robert Charles  
; APPLICANT: Guterman, Sonia Kosow  
; APPLICANT: Roberts, Bruce Lindsay  
; APPLICANT: Markland, William  
; APPLICANT: Ley, Arthur Charles  
; APPLICANT: Kent, Rachel Baribault  
; TITLE OF INVENTION: Directed Evolution of No. 5223409el  
; TITLE OF INVENTION: Binding Proteins  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Browdy and Neimark  
; STREET: 419 Seventh Street, N.W.  
; CITY: Washington,  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 4.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/664,989B  
; FILING DATE: 19910301  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US89/03731  
; FILING DATE: 01-SEP-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/487,063  
; FILING DATE: 02-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/240,160  
; FILING DATE: 02-SEP-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cooper, Iver P.  
; REGISTRATION NUMBER: 28005  
; REFERENCE/DOCKET NUMBER: LADNER 7  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-664-989B-11

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-07-664-989B-11 ..

Align seg 1/1 to: US-07-664-989B-11 from: 1 to: 15

2846 TCTCTGGGTCTTCCAGTTTG 2826

|||||  
9 SerLeuGlySerSerSerLeu 15

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-07-664-989B-18

seq\_documentation\_block:  
; Sequence 18, Application US/07664989B  
; Patent No. 5223409  
; GENERAL INFORMATION:  
; APPLICANT: Ladner, Robert Charles  
; APPLICANT: Guterman, Sonia Kosow  
; APPLICANT: Roberts, Bruce Lindsay  
; APPLICANT: Markland, William  
; APPLICANT: Ley, Arthur Charles  
; APPLICANT: Kent, Rachel Baribault  
; TITLE OF INVENTION: Directed Evolution of No. 5223409el  
; TITLE OF INVENTION: Binding Proteins  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Browdy and Neimark  
; STREET: 419 Seventh Street, N.W.  
; CITY: Washington,  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 4.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/664,989B  
; FILING DATE: 19910301  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US89/03731  
; FILING DATE: 01-SEP-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/487,063  
; FILING DATE: 02-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/240,160  
; FILING DATE: 02-SEP-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cooper, Iver P.  
; REGISTRATION NUMBER: 28005  
; REFERENCE/DOCKET NUMBER: LADNER 7  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-664-989B-18

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-07-664-989B-18 ..

Align seg 1/1 to: US-07-664-989B-18 from: 1 to: 15

2846 TCCTGGGCTCCAGTTG 2826

|||||

9 SerLeuGlySerSerLeu 15

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pap:US-08-268-251-24

seq\_documentation\_block:  
; Sequence 24, Application US/08268251  
; Patent No. 5585475  
; GENERAL INFORMATION:

; APPLICANT: Jamieson, Gordon A  
; APPLICANT: Dedman, John R  
; APPLICANT: Kaetzel, Marcia A  
; TITLE OF INVENTION: Calmodulin-Binding Peptides  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/268,251  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/831,219  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Grant D  
; REGISTRATION NUMBER: 31,259  
; REFERENCE/DOCKET NUMBER: 272,001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2706  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-268-251-24

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-08-268-251-24 ..

Align seg 1/1 to: US-08-268-251-24 from: 1 to: 15

830 TTGAGGTATTAAATCATCT 810

|||||

8 LeuArgLeuLeuLysSer 14

seq\_name: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pap:PCT-US93-01112-24

seq\_documentation\_block:

; Sequence 24, Application PC/TUS9301112

; GENERAL INFORMATION:

; APPLICANT: Jamieson, Gordon A  
; APPLICANT: Dedman, John R  
; APPLICANT: Kaetzel, Marcia A  
; TITLE OF INVENTION: Calmodulin-Binding Peptides  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/01112  
; FILING DATE: 19930208  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/831,219  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Grant D  
; REGISTRATION NUMBER: 31,259  
; REFERENCE/DOCKET NUMBER: 272,001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2706  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: AMINO ACID

```
; ; STRANDEDNESS: single
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: peptide
PCT-US93-01112-24

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x PCT-US93-01112-24
Align seg 1/1 to: PCT-US93-01112-24 from: 1 to: 15

      830 TTGAGGTTATTAAATCATCT 810
      |||||||
      8 LeuArgLeuLeuLysSer 14

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-268-251-54

seq_documentation_block:
; Sequence 54, Application US/08268251
; Patent No. 5585475
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/268,251
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-268-251-54

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
```

```
US-09-202-054-2/rev x US-08-268-251-54
Align seg 1/1 to: US-08-268-251-54 from: 1 to: 23

      830 TTGAGGTTATTAAATCATCT 810
      |||||||
      10 LeuArgLeuLeuLysSer 16

seq_name: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-01112-54

seq_documentation_block:
; Sequence 54, Application PC/TUS9301112
; GENERAL INFORMATION:
; APPLICANT: Jamieson, Gordon A
; APPLICANT: Dedman, John R
; APPLICANT: Kaetzel, Marcia A
; TITLE OF INVENTION: Calmodulin-Binding Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01112
; FILING DATE: 19930208
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-01112-54

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x PCT-US93-01112-54
Align seg 1/1 to: PCT-US93-01112-54 from: 1 to: 23

      830 TTGAGGTTATTAAATCATCT 810
      |||||||
      10 LeuArgLeuLeuLysSer 16

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-486-013-17

seq_documentation_block:
; Sequence 17, Application US/08486013
; Patent No. 5731149
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Seisted, Michael E.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486.013
; FILING DATE:
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/342,268
; FILING DATE: 18-NOV-1994
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 1206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-486-013-17

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-202-054-2/rev x US-08-486-013-17  ..

  Align seg 1/1 to: US-08-486-013-17 from: 1 to: 45

1912 AAGAGATGTCATTGCGTTCA 1892
|||||
6 LysArgCysHisCysArgSer 12

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-482-279-17

seq_documentation_block:
; Sequence 17, Application US/08482279
; Patent No. 5840498
; GENERAL INFORMATION:
; APPLICANT: Seisted, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; TITLE OF INVENTION: of Their Use
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
;
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,279
; FILING DATE:
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/342,268
; FILING DATE: 18-NOV-1994
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 1206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-482-279-17

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-202-054-2/rev x US-08-482-279-17  ..

  Align seg 1/1 to: US-08-482-279-17 from: 1 to: 45

1912 AAGAGATGTCATTGCGTTCA 1892
|||||
6 LysArgCysHisCysArgSer 12

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-342-268-17

seq_documentation_block:
; Sequence 17, Application US/08342268
; Patent No. 5844072
; GENERAL INFORMATION:
; APPLICANT: Seisted, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; TITLE OF INVENTION: of Their Use
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/342,268
; FILING DATE: 18-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 1206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-342-268-17

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x US-08-342-268-17 ..
Align seg 1/1 to: US-08-342-268-17 from: 1 to: 45

1912 AGAGATGTCATTGTCGTCA 1892
|||||
6 LysArgCysHisCysArgSer 12

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pap:US-09-015-968-17

seq_documentation_block:
; Sequence 17, Application US/09015968
; Patent No. 6057425
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; TITLE OF INVENTION: of Their Use
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,968
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,279
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/342,268
; FILING DATE: 18-NOV-1994
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 3003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-09-015-968-17

alignment_scores:
  Quality: 7.00      Length: 7
  Ratio: 1.000      Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-202-054-2/rev x US-09-015-968-17 ..
Align seg 1/1 to: US-09-015-968-17 from: 1 to: 45

1912 AGAGATGTCATTGTCGTCA 1892
|||||
6 LysArgCysHisCysArgSer 12

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-09-397-386-17

seq_documentation_block:
; Sequence 17, Application US/09397386
; Patent No. 6300470
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; TITLE OF INVENTION: of Their Use
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/397,386
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,968
; FILING DATE:
; APPLICATION NUMBER: US 08/482,279
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/342,268
; FILING DATE: 18-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: US 07/889,020  
FILING DATE: 26-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-UC 3003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-397-386-17

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-09-397-386-17 ..

Align seg 1/1 to: US-09-397-386-17 from: 1 to: 45

1912 AAGAGATGTCATGCTGTTCA 1892

|||||

6 LysArgCysHisCysArgSer 12

seq\_name: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:US92-08558-9

seq\_documentation\_block:

Sequence 9, Application PC/TUS9208558

GENERAL INFORMATION:

APPLICANT: Cornell Research Foundation, Inc.

TITLE OF INVENTION: MOLECULAR CLONING AND TRANSFORMATION OF CYCLODIENE RESISTANCE

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Yahwak & Associates

STREET: 25 Skytop Drive

CITY: Trumbull

STATE: Connecticut

COUNTRY: USA

ZIP: 06611

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: MS-DOS

SOFTWARE: Microsoft Word 4.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/08558

FILING DATE: 19921002

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/770,881

FILING DATE: October 4th 1991

ATTORNEY/AGENT INFORMATION:

NAME: George M. Yahwak

REGISTRATION NUMBER: 26,824

REFERENCE/DOCKET NUMBER: CRF D-1052

TELECOMMUNICATION INFORMATION:

TELEPHONE: (203)268-1951

TELEFAX: (203)268-1951

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 53 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US92-08558-9

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x PCT-US92-08558-9 ..

Align seg 1/1 to: PCT-US92-08558-9 from: 1 to: 53

968 AAAACTTTTAATTCTGCAGC 948

|||||

36 LysThrPheAsnSerValSer 42

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-152-721B-28

seq\_documentation\_block:

Sequence 28, Application US/08152721B

Patent No. 5962315

GENERAL INFORMATION:

APPLICANT: Livingston, David M.

APPLICANT: Ewen, Mark E.

TITLE OF INVENTION: DNA Encoding p107 Tumor Suppressor and

TITLE OF INVENTION: Related Polypeptides

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHOATE, HALL & STEWART

STREET: 53 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2891

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/152,721B

FILING DATE: 15-NOV-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Pasternack Esq., Sam

REGISTRATION NUMBER: 29,576

REFERENCE/DOCKET NUMBER: 181411-011DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-5000

TELEFAX: (617) 248-4000

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 71 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

IMMEDIATE SOURCE:

CLONE: RB region 6

US-08-152-721B-28

alignment\_scores:

Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2/rev x US-08-152-721B-28 ..

Align seg 1/1 to: US-08-152-721B-28 from: 1 to: 71



1152 ATAGACCTGAAGTTCAAAATT 1132  
|||||  
43 IleAspLeuLysPheLysIle 49  
seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-09-130-339-4  
seq\_documentation\_block:  
; Sequence 4, Application US/09130339  
; Patent No. 6358701  
; GENERAL INFORMATION:  
; APPLICANT: Cully, Doris F.  
; APPLICANT: Pares, Philip S.  
; APPLICANT: Warmke, Jeffrey W.  
; APPLICANT: Etter, Adrian  
; APPLICANT: Cohen, Charles J.  
; APPLICANT: Brochu, Richard M.  
; TITLE OF INVENTION: DNA MOLECULES ENCODING CTENOCEPHALIDES  
; TITLE OF INVENTION: FELIS GLUTAMATE GATED CHLORIDE CHANNELS  
; FILE REFERENCE: 20029  
; CURRENT APPLICATION NUMBER: US/09/130,339  
; CURRENT FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: ctenocephalides felis  
US-09-130-339-4  
alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2/rev x US-09-130-339-4 ..  
Align seg 1/1 to: US-09-130-339-4 from: 1 to: 78  
2974 TCTTGCATACCTGCTGTCA 2954  
|||||  
15 SerLeuHisThrCysLeuSer 21  
seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-710-749-13  
seq\_documentation\_block:  
; Sequence 13, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Brilles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid

; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-13  
alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-202-054-2 x US-08-710-749-13 ..  
Align seg 1/1 to: US-08-710-749-13 from: 1 to: 99  
2029 CTGCTAAATTTAGAGGAATTA 2049  
|||||  
32 LeuLeuLysLeuGlulLeu 38  
seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-710-749-14  
seq\_documentation\_block:  
; Sequence 14, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Brilles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid

US-08-710-749-14

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x US-08-710-749-14 ..

Align seg 1/1 to: US-08-710-749-14 from: 1 to: 99

2029 CTGCTAAATAGAGGAATTA 2049

|||||  
32 LeuLeuLysLeuGluGluLeu 38

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-710-749-16

seq\_documentation\_block:  
; Sequence 16, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 16:  
; LENGTH: 99 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
US-08-710-749-16

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-08-710-749-16 ..

Align seg 1/1 to: US-08-710-749-16 from: 1 to: 99

2029 CTGCTAAATAGAGGAATTA 2049

|||||  
32 LeuLeuLysLeuGluGluLeu 38

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-340-203A-12

seq\_documentation\_block:  
; Sequence 12, Application US/08340203A  
; Patent No. 5756668  
; GENERAL INFORMATION:  
; APPLICANT: Baylin, Stephen B.  
; APPLICANT: Wales, Michele M.  
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/340,203A  
; FILING DATE: 15-NOV-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A.  
; REGISTRATION NUMBER: P-38,347  
; REFERENCE/DOCKET NUMBER: 07265/039001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 455-5100  
; TELEFAX: (619) 455-5110  
; INFORMATION FOR SEQ ID NO: 12:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-340-203A-12

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-202-054-2 x US-08-340-203A-12 ..

Align seg 1/1 to: US-08-340-203A-12 from: 1 to: 102

1713 GCTGAGATATTGGACTTCTC 1733

|||||  
65 AlaGluIlePheGlyLeuLeu 71

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-452-427-12

seq\_documentation\_block:  
; Sequence 12, Application US/08452427  
; Patent No. 5922590  
; GENERAL INFORMATION:  
; APPLICANT: Baylin, Stephen B.  
; APPLICANT: Wales, Michele M.  
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/452,427  
; FILING DATE: 26-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/340,203  
; FILING DATE: 15-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A.  
; REGISTRATION NUMBER: P-38,347  
; REFERENCE/DOCKET NUMBER: 07265/039001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 455-5100  
; TELEFAX: (619) 455-5110  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-452-427-12

alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-202-054-2 x US-08-452-427-12 ..  
Align seg 1/1 to: US-08-452-427-12 from: 1 to: 102

1713 GGTGAGATATTGGACTTCTC 1733  
|||||  
65 AlaGluIlePheGlyLeuLeu 71

seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-710-749-8

seq\_documentation\_block:  
; Sequence 8, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749

; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454312-2074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: n/a  
; TOPOLOGY: linear  
; MOLECULE TYPE: amino acid  
; US-08-710-749-8  
  
alignment\_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
  
alignment\_block:  
US-09-202-054-2 x US-08-710-749-8 ..  
Align seg 1/1 to: US-08-710-749-8 from: 1 to: 102  
  
2029 CTGCTAAATTTAGAGGAATTA 2049  
|||||  
33 LeuLeuLysLeuGluGluLeu 39

